

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 20:35:55 ; Search time 53 Seconds  
(without alignments)  
90.628 Million cell updates/sec

Title: US-09-982-259b-7

Perfect score: 89  
Sequence: 1 GMTFRAQEGAFITGAAC 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_29Jan04.\*  
1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001s.\*  
5: Geneseq2002s.\*  
6: Geneseq2003as.\*  
7: Geneseq2003bs.\*  
8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	100.0	17	5	ABB78091 Amino aci
2	72	80.9	14	6	ABU61976 Borellia
3	72	80.9	14	6	ABU62398 Lyme dise
4	72	80.9	322	2	AAY19985 B. burgdo
5	72	80.9	339	2	AAY19985 B. burgdo
6	72	80.9	339	2	AAY19984 B. burgdo
7	69	77.5	340	2	AAR31013 P39-alpha
8	44.5	50.0	604	4	ABG22311 Novel hum
9	43.5	48.9	4924	4	AAB70968 S. spinos
10	43.5	48.9	4928	4	AAY39300 Spnd a po
11	43.5	48.9	4933	6	ABP57681 Saccharop
12	43	48.3	616	5	ABP57681 Saccharop
13	42	47.2	341	2	ABP57681 Saccharop
14	42	47.2	341	2	AAR33280 Herbicida
15	41	46.1	130	5	AAR33280 P39-beta
16	41	46.1	130	5	AAR33280 P39-beta
17	41	46.1	247	3	ABG76529 HCV El an
18	41	46.1	1366	4	ABG76529 HCV El an
19	40	44.9	59	4	ABU63665 Bacillus m
20	40	44.9	59	4	ABU63665 Bacillus m
21	40	44.9	117	6	AEO10127 Streptomy
22	40	44.9	196	5	AEO10127 Streptomy
23	40	44.9	200	5	ADA54805 Human pol
24	40	44.9	328	6	ABP30378 Streptoco
25	40	44.9	339	2	ABP30378 Streptoco
26	40	44.9	339	2	AAY00049 Enterococ
27	40	44.9	339	2	ABP43268 E faecali

26	40	44.9	339	6	ABU088296	Abu88296 E. faecal
27	40	44.9	339	6	ABU13547	Abu13547 Enterococ
28	40	44.9	361	2	AAY00048	Aay00048 Enterococ
29	40	44.9	361	5	ABP43267	Abp43267 E faecali
30	40	44.9	361	6	ABU088295	Abu088295 E. faecal
31	40	44.9	361	6	ABU13546	Abu13546 Enterococ
32	40	44.9	537	4	AAB96626	Aab96626 Putative
33	40	44.9	604	4	AAW40228	Aaw40228 Human pol
34	40	44.9	604	6	ABR57251	AbR57251 Human mit
35	40	44.9	621	4	AAU33270	Aau33270 Novel hum
36	40	44.9	624	4	AAW42014	Aaw42014 Human pol
37	40	44.9	625	4	ABG21889	Abg21889 Novel hum
38	39.5	44.4	1891	2	AAW22610	Aaw22610 Platenoli
39	39.5	44.4	1891	2	AAW23720	Aaw23720 Platenoli
40	39	43.8	147	5	ABP11171	Abp11171 Human ORF
41	39	43.8	317	4	AAU41886	Aau41886 Propionib
42	39	43.8	317	6	ABM38405	Abm38405 Propionib
43	39	43.8	494	6	ABU39165	Abu39165 Protein e
44	39	43.8	561	4	ABG04039	Abg04039 Novel hum
45	39	43.8	1168	5	ABP26042	Abp26042 Streptoco

## ALIGNMENTS

RESULT 1  
ABB78091  
ID ABB78091 standard; peptide; 17 AA.  
XX  
AC ABB78091;  
XX  
DT 06-AUG-2003 (revised)  
DT 05-NOV-2002 (first entry)  
XX  
DE Amino acid sequence of a Borellia burgdorferii epitope.  
XX  
KW Epitope; polyethylene glycol; PEG; vaccine; Lyme disease.  
XX  
OS Synthetic.  
OS Borellia burgdorferi.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 15  
FT Modified-site 16 /note= "beta alanine"  
FT Modified-site 16 /note= "beta alanine\  
XX  
FN WO200234117-A2.  
XX  
PD 02-MAY-2002.  
XX  
PF 22-OCT-2001; 2001WO-US046723.  
XX  
PR 24-OCT-2000; 2000US-0242819P.  
PR 17-OCT-2001; 2001US-00982264.  
XX  
PA (UNVE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.  
XX  
PI Qiu B, Stein S, Zhang G, Sigal L, Brunner M, Katz M;  
XX WPI; 2002-618962/66.  
XX  
PT An immunological test kit for detecting analytes, comprising a  
PT composition of matter with an immunologically reactive substance e.g.  
PT Borellia burgdorferi epitope polypeptide linked to an immunologically  
PT invisible carrier.  
XX  
CC Claim 51; Page 25; 36pp; English.  
XX  
CC ABB78085-91 represent Borellia burgdorferi epitope polypeptides. The  
CC peptides are conjugated in multiple copies to an immunologically  
CC invisible carrier. This carrier is preferably polyethylene glycol (PEG)  
CC copolymer of a structure given in the specification. The conjugated

CC peptides are used as vaccines, for immunising against or treating Lyme disease. They may also be used in an immunological test kit (such as enzyme linked immunosorbent assay (ELISA) kit, or immuno-capillary kit).  
 CC ('Updated on 06-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 17 AA;

Query Match 100.0%; Score 89; DB 5; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-09; Mismatches 0; Indels 0; Gaps 0;  
 Matches 17; Conservative 0;

QY 1 GMTFRAQEGAFLTGAAC 17  
 |||||  
 Db 1 GMTFRAQEGAFLTGAAC 17

RESULT 2  
 ABU61976  
 ID ABU61976 standard; peptide; 14 AA.

XX AC ABU61976;

XX DT 25-AUG-2003 (first entry)

XX DE Borellia burgdorferii epitope P39.

XX KW Epitope; vaccine; polyethylene glycol; PEG; ELISA; Lyme disease;  
 XX enzyme linked immunosorbent assay; immuno-capillary kit.

XX OS Borellia burgdorferii.

XX FH Key Location/Qualifiers  
 XX Modified-site 14

XX FT /label= OTHER

XX FT /note= "Gly is linked to a (beta A) (beta A) C moiety  
 XX (not defined)"

XX PN US2003040127-A1.

XX PD 27-FEB-2003.

XX PF 17-OCT-2001; 2001US-00982287.

XX PR 24-OCT-2000; 2000US-0242819P.

XX PA (QIUB/) QIU B.

XX PA (STEI/) STEIN S.

XX PA (ZHAN/) ZHANG G.

XX PA (SIGA/) SIGAL L.

XX PA (BRUN/) BRUNNER M.

XX PA (KATZ/) KATZ M.

XX PI Qiu B, Stein S, Zhang G, Sigal L, Brunner M, Katz M;

XX DR WPI; 2003-503407/47.

XX Novel composition of matter for use in immunological test kit, comprises  
 PT a first immunologically reactive substance connected to a second  
 PT immunologically reactive substance by an immunologically invisible  
 PT carrier.

XX PS Disclosure; Page 5; 13pp; English.

XX The invention relates to a composition of matter (I), comprising a first  
 CC immunologically reactive substance (Iia) connected to a second  
 CC immunologically reactive substance (Iib) by an immunologically invisible  
 CC carrier (polyethylene glycol, PEG). The composition is useful in an  
 CC immunological test kit e.g. an enzyme linked immunosorbent assay (ELISA)  
 CC kit and immuno-capillary kit. The composition is also useful as vaccine  
 CC and as an in vivo therapeutic and in immuno-chromatography, for isolating  
 CC and purifying desired antibody in quantity, and for detecting low levels  
 CC of antibody in clinical samples. The present sequence is an epitope (used  
 CC in the composition of the invention) from Borellia burgdorferii (the

CC causative agent of Lyme disease)  
 XX  
 SQ Sequence 14 AA;

Query Match 80.9%; Score 72; DB 6; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 5e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFLTG 14  
 |||||  
 Db 1 GMTFRAQEGAFLTG 14

RESULT 3  
 ABU62398

ID ABU62398 standard; peptide; 14 AA.

XX AC ABU62398;

XX DT 01-SEP-2003 (first entry)

XX DE Lyme disease spirochete epitope polypeptide P39.

XX KW Lyme disease spirochete; Lyme disease; immunological test kit; epitope;  
 XX invisible carrier; P39.

XX OS Borrelia burgdorferi.

XX PN US2003040126-A1.

XX PD 27-FEB-2003.

XX PF 17-OCT-2001; 2001US-00982265.

XX PR 24-OCT-2000; 2000US-0242819P.

XX PA (QIUB/) QIU B.

XX PA (ZHAN/) ZHANG G.

XX PA (STEI/) STEIN S.

XX PA (SIGA/) SIGAL L H.

XX PA (BRUN/) BRUNNER M.

XX PA (KATZ/) KATZ M.

XX PI Qiu B, Zhang G, Stein S, Sigal LH, Brunner M, Katz M;

XX DR WPI; 2003-503406/47.

XX Immunological test kit comprises immunologically reactive substance  
 connected to immunologically invisible carrier.

XX PS Disclosure; Page 1; 13pp; English.

XX The invention relates to an immunological test kit including a  
 CC composition of matter comprising an immunologically reactive substance  
 CC connected to an immunologically invisible carrier. The invention is  
 CC useful as an immunological test kit and can be used to detect disease  
 CC e.g. Lyme disease. The invention allows disease diagnosis even at an  
 CC early stage which is much more sensitive. The present sequence represents  
 CC the amino acid sequence of the Lyme disease spirochete epitope  
 CC polypeptide P39

XX SQ Sequence 14 AA;

Query Match 80.9%; Score 72; DB 6; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 5e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFLTG 14  
 |||||  
 Db 1 GMTFRAQEGAFLTG 14

RESULT 4

```

XX 14-JUL-1998.
XX
XX
XX 01-MAR-1995; 95US-00396957.
XX
XX 05-MAR-1990; 90US-00487716.
XX 05-MAR-1991; 91US-00684731.
XX 19-FEB-1993; 93US-00020245.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Schwan TG, Simpson WJ;
XX
XX WPI; 1998-413001/35.
XX N-PSDB; AAV35685, AAV35686.
XX
XX New antigenic proteins from Borrelia burgdorferi useful in vaccines - for
XX raising antibodies or for diagnostic detection of specific antibodies.
XX
XX Claim 1; Col 29-32; 33pp; English.
XX
XX This represents the Borrelia burgdorferi antigenic protein p39 alpha
XX which is immunoreactive with mammalian Lyme borreliosis serum. The
XX Borrelia antigenic proteins p39 alpha and p39 beta are used in vaccines
XX to protect against Lyme disease, as assay reagents to detect specific
XX antibodies in the serum (diagnostic of Lyme disease), and to raise
XX antibodies, either for diagnosis (by detecting the corresponding antigen)
XX or in screening agents for ability to inhibit expression of the proteins
XX
XX SQ Sequence 339 AA;
XX
XX Query Match 80.9%; Score 72; DB 2; Length 339;
XX Best Local Similarity 100.0%; Pred. NO. 0.00017;
XX Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 GMTFRAQEGAFLLG 14
XX |||||
XX 129 GMTFRAQEGAFLLG 142
XX
XX RESULT 6
XX AAY19984
XX ID AAY19984 standard; protein; 339 AA.
XX AC AAY19984;
XX
XX 19-JUL-1999 (first entry)
XX
XX B. burgdorferi antigenic protein, f541.aa.
XX
XX Antigenic protein; vaccine; Lyme disease; infection; detection.
XX
XX Borrelia burgdorferi.
XX
XX W09859071-AL.
XX
XX 30-DEC-1998.
XX
XX 18-JUN-1998; 98WO-US012718.
XX
XX 20-JUN-1997; 97US-0050359P.
XX 22-JUL-1997; 97US-0053344P.
XX 22-JUL-1997; 97US-0053377P.
XX 03-SEP-1997; 97US-0057483P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX (MEDI-) MEDIMUNE INC.
XX
XX Choi GH, Erwin AL, Hanson MS, Lathigra R;
XX
XX WPI; 1999-189980/16.
XX N-PSDB; AAX61681.
XX

```

PT New isolated Borrelia burgdorferi nucleic acids - used to develop  
 PT products for the diagnosis, prevention and treatment of diseases caused  
 PT by Borrelia, particularly Lyme disease.

XX Claim 12; Page 149-150; 275pp; English.

CC This sequence represents a Borrelia burgdorferi (Bb) protein of the  
 CC invention, which is suitable for use in a vaccine. The Bb polypeptides  
 CC can be used in vaccines for eliciting protective antibodies to members of  
 CC the Borrelia genus, particularly for the use against Lyme disease in  
 CC humans and animals. They can be used for preventing or attenuating an  
 CC infection caused by a member of the Borrelia genus. The products can also  
 CC be used for detection of members of the Borrelia genus

XX Sequence 339 AA;

Query Match 80.9%; Score 72; DB 2; Length 339;  
 Best Local Similarity 100.0%; Pred. No. 0.00017;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFITG 14  
 |||||:|||||  
 Db 129 GMTFRAQEGAFITG 142

# RESULT 7

AA331013  
 ID AAR31013 standard; protein; 340 AA.

XX AAR31013;

XX 25-MAR-2003 (revised)

XX 17-DEC-2001 (revised)

XX 11-MAY-1993 (first entry)

XX P39-alpha.

XX Polymerase chain reaction; PCR; primer; open reading frame; ORF; P39;  
 KW antigen; 39-alpha; 39-beta; serum; human; Lyme patient; p39-beta;  
 KW p39-alpha; promoter; -10; -35; ribosomal binding site; RBS.

XX Borrelia burgdorferi.

XX USN7664731-N.

XX 01-DEC-1992.

XX 05-MAR-1991; 91US-00664731.

XX 05-MAR-1990; 90US-00487716.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICE.

XX Simpson WJ, Schwan TG;

XX W2; 1993-017799/02.

XX N-PSDB; AAQ34672.

PT New Borrelia burgdorferi protein and DNA isolates - used in detection of  
 PT Lyme borreliosis infection and in prodn. of antibodies and vaccines.

XX Disclosure; Page 36-39; 69pp; English.

CC The sequences given in AAR31013 and AAR33280 represent the B. burgdorferi  
 CC proteins P39-alpha and P39-beta. The DNA encoding these proteins was  
 CC isolated by polymerase chain reaction (PCR) using the primer sequences  
 CC given in AAQ34673-76. This DNA contained two open reading frames (ORF).  
 CC It was therefore concluded that the P39 antigen which has been previously  
 CC described is not one protein but two, 39-alpha and 39-beta. The P39  
 CC signal appears to be enhanced when both genes are present. Gene 1 encodes  
 CC a 339 amino acid protein with a calculated molecular weight of 36.926 kD.  
 CC The protein encoded by this gene which is reactive with serum from human  
 CC Lyme patients. The ORF of gene 2 has been designated p39-beta. This

CC genes' ORF starts 116 nucleotides downstream of p39-alpha and encodes a  
 CC protein of 341 amino acids (37,506 kD). The promoter located 5' of p39-  
 CC alpha appears to be present with classic -10 and -35 regions whereas p39-  
 CC beta lacks a recognisable promoter sequence. Both genes have putative  
 CC ribosomal binding sites immediately 5' to the start codons and each is  
 CC terminated with a TAA codon. Comparing the gene sequences of p39-alpha  
 CC and p39-beta indicates that these genes have 62% similarity. (Note:  
 CC Revised entry submitted to correct the patent number format of US  
 CC Government-owned NTRIS applications to prevent clashes with ongoing US  
 CC granted patent numbers. For further information please visit the Derwent  
 CC web site at [www.derwent.com/dwpi/updates/ntris\\_us.html](http://www.derwent.com/dwpi/updates/ntris_us.html).) (Updated on 25-  
 CC MAR-2003 to correct PF field.)

XX Sequence 340 AA;

Query Match 77.5%; Score 69; DB 2; Length 340;  
 Best Local Similarity 92.9%; Pred. No. 0.00061;  
 Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFITG 14  
 |||||:|||||  
 Db 130 GMTFRAQEGAFITG 143

# RESULT 8

ABG22311

ID ABG22311 standard; protein; 604 AA.

XX ABG22311;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #22302.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS86498.

XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.

XX Claim 20; SEQ ID NO 52670; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 604 AA;

Query Match 50.0%; Score 44.5; DB 4; Length 604;  
 Best Local Similarity 58.8%; Pred. No. 32;  
 Matches 10; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 2 MTFRAQEG-APLTGAAC 17

DB 546 MKFKASEGPAFFPGAMC 562

# RESULT 9

AA070968  
 ID AAB70968 standard; protein; 4924 AA.

XX AAB70968;

XX 28-AUG-2001 (first entry)

XX S. spinosa protein fragment encoded by ORF21, SEQ ID 48.

XX Forosamine; trimethylrhamsone; polyketide synthase; biosynthesis;  
 KW spinosyn; polyketide aglycone; transgenic plant; insect resistance;  
 KW macrolide; insecticidal; polyketide synthase.

OS Saccharopolyspora spinosa.

XX DE19957268-A1.

XX 08-MAR-2001.

XX 29-NOV-1999; 99DE-01057268.

XX 27-AUG-1999; 99DE-01040596.

XX (FARB ) BAYER AG.

XX Eberz G, Moehrle V, Froede R, Velten R, Salas JA;

XX WPI; 2001-267102/28.

XX N-PSDB; AAF88338.

XX New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for  
 PT recombinant production of insecticidal spinosyns and their derivatives.

XX Claim 57; Page 264-284; 354pp; German.

XX This invention describes a novel method nucleic acid (I) and its encoded  
 CC polypeptide (II) containing at least one region that encodes an enzymatic  
 CC activity involved in biosynthesis of spinosyns. (I) are used (i) to  
 CC identify, inactivate or modulate genes involved in the biosynthesis of  
 CC (II); (ii) to generate a library of polyketide synthases; (iii) for  
 CC adding forosamine or trimethylrhamsone to a spinosyn or polyketide  
 CC aglycone; and (iv) for recombinant production of the corresponding  
 CC enzymes, which are used for production of (II), their precursors or  
 CC derivatives, including production of transgenic plants that express (II)  
 CC and thus have increased resistance to insects. (I) are also useful as  
 CC markers for sequencing of the Saccharopolyspora spinosa genome. (II) are  
 CC macrolides with insecticidal, but not antibacterial, activity, and can  
 CC also be used to raise specific antibodies, useful for identifying  
 CC expression clones in a gene bank. Cells transformed with (I) may produce  
 CC (II) at significantly increased levels or produce new derivatives of

CC (II). This sequence represents an S. spinosa polyketide synthase  
 XX Sequence 4924 AA;

Query Match 48.9%; Score 43.5; DB 4; Length 4924;  
 Best Local Similarity 66.7%; Pred. No. 5e+02;  
 Matches 10; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 GMTFRAQEGAPLTGA 15

DB 3379 GSTY-VREGAPLTGA 3392

# RESULT 10

AA070968

ID AAY39300 standard; protein; 4928 AA.

XX AAY39300;

XX 01-DEC-1999 (first entry)

XX SpnD a polyketide synthase.

XX Spinosyn biosynthetic enzyme; open reading frame; ORF;  
 KW insecticidal macrolides; arachnid; nematode; insect; polyketide;  
 KW polyketide synthase; PKS; extender module; initiator module;  
 KW acyl transferase domain; AT; acyl carrier protein; ACP;  
 KW beta-ketosynthase domain; KS; KR; dehydratase domain; DH;  
 KW enoyl reductase domain; ER; beta-ketoreductase; insecticide.

XX Saccharopolyspora spinosa.

XX Key Location/Qualifiers

PH Domain

FT 1..424

FT /label= KS5

FT /note= "Beta-ketosynthase domain: part of extender module

5"

FT 539..866

FT /label= AT5

FT /note= "Acyl transferase domain: part of extender module

5"

FT 893..1078

FT /label= DHS

FT /note= "Dehydratase domain: part of extender module 5"

FT 1384..1565

FT /label= KRS

FT /note= "Beta-ketoreductase domain: part of extender

module 5"

FT 1645..1726

FT /label= ACP5

FT /note= "Acyl carrier protein domain: part of extender

module 5"

FT 1748..2172

FT /label= KS6

FT /note= "Beta-ketosynthase domain: part of extender module

6"

FT 2283..2613

FT /label= AT6

FT /note= "Acyl transferase domain: part of extender module

6"

FT 2916..3095

FT /label= KR6

FT /note= "Beta-ketoreductase domain: part of extender

module 6"

FT 3188..3269

FT /label= ACP6

FT /note= "Acyl carrier protein domain: part of extender

module 6"

FT 3291..3713

FT /label= KS7

FT /note= "Beta-ketosynthase domain: part of extender module

7"

FT 3825..4153

FT Domain



XX KW Herbicidal; plant; agriculture; herbicide.  
 XX OS Arabidopsis thaliana.  
 XX FN WO200210210-A2.  
 XX PD 07-FEB-2002.  
 XX PF 28-AUG-2001; 2001WO-EP009892.  
 XX PR 28-AUG-2001; 2001WO-EP009892.  
 XX PA (FARB ) BAYER AG.  
 XX PI Tietjen K, Weidner M;  
 XX DR WPI; 2002-269010/31.  
 XX PT Identifying plant target proteins for herbicidally active compounds,  
 PT comprising aligning and comparing nucleic acid or amino acid sequences  
 PT from plant with nucleic acid or amino acid sequences from non-plant  
 PT organisms.  
 XX PS Claim 5; SEQ ID NO 311; 261pp + Sequence Listing; English.  
 XX CC The invention relates to identifying target proteins (ABB90790-ABB94016)  
 CC for herbicidally active compounds, comprising aligning and comparing  
 CC nucleic acid or amino acid sequences from plant with nucleic acid or  
 CC amino acid sequences from non-plant organisms using suitable search  
 CC parameters, where plant sequences having an E-value greater by a factor  
 CC of 3 than the E-value of most similar non-plant sequences are selected.  
 CC The polypeptides or nucleic acids encoding them are useful for  
 CC identifying modulators. The identified modulators are useful as  
 CC herbicides  
 XX SQ Sequence 616 AA;  
 Query Match 48.3%; Score 43; DB 5; Length 616;  
 Best Local Similarity 57.1%; Pred. No. 61;  
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 TFRQEGAFITGAA 16  
 |||::|||::|||  
 DB 3: TFRSQENLYDGA 44  
 RESULT 13  
 AAR33280  
 ID AAR33280 standard; protein; 341 AA.  
 AC AAR33280;  
 XX AC  
 XX DT 25-MAR-2003 (revised)  
 DT 17-DEC-2001 (revised)  
 DT 11-MAY-1993 (first entry)  
 XX P39-beta.  
 XX DE  
 XX KW Polymerase chain reaction; PCR; primer; open reading frame; ORF; P39;  
 KW antigen; 39-alpha; 39-beta; serum; human; Lyme patient; p39-beta;  
 KW p39-alpha; promoter; -10; -35; ribosomal binding site; RBS.  
 XX OS Borrelia burgdorferi.  
 XX PN USN7664731-N.  
 XX PF 01-DEC-1992.  
 XX PD 05-MAR-1991; 91US-00664731.  
 XX PF 05-MAR-1991; 91US-00664731.  
 XX PR 05-MAR-1990; 90US-00487716.  
 XX PR 05-MAR-1990; 90US-00487716.  
 (USSH ) US DEPT HEALTH & HUMAN SERVICE.  
 Simpson WJ, Schwan TG;  
 WPI; 1993-017799/02.  
 N-PSDB; AAQ34672.  
 New Borrelia burgdorferi protein and DNA isolates - used in detection of  
 Lyme borreliosis infection and in prodn. of antibodies and vaccines.  
 Disclosure; Page 36-39; 69pp; English.  
 The sequences given in AAR31013 and AAR33280 represent the B. burgdorferi  
 proteins p39-alpha and p39-beta. The DNA encoding these proteins was  
 isolated by polymerase chain reaction (PCR) using the primer sequences  
 given in AAQ34673-76. This DNA contained two open reading frames (ORF).  
 It was therefore concluded that the p39 antigen which has been previously  
 described is not one protein but two, 39-alpha and 39-beta. The p39  
 signal appears to be enhanced when both genes are present. Gene 1 encodes  
 a 339 amino acid protein with a calculated molecular weight of 36,926 kD.  
 The protein encoded by this gene which is reactive with serum from human  
 Lyme patients. The ORF of gene 2 has been designated p39-beta. This  
 genes' ORF starts 116 nucleotides downstream of p39-alpha and encodes a  
 protein of 341 amino acids (37,506 kD). The promoter located 5' of p39-  
 alpha appears to be present with classic -10 and -35 regions whereas p39-  
 beta lacks a recognisable promoter sequence. Both genes have putative  
 ribosomal binding sites immediately 5' to the start codons and each is  
 terminated with a TAA codon. Comparing the gene sequences of p39-alpha  
 and p39-beta indicates that these genes have 62% similarity. (Note:  
 Revised entry submitted to correct the patent number format of US  
 Government-owned NTIS applications to prevent clashes with ongoing US  
 granted patent numbers. For further information please visit the Derwent  
 web site at [www.derwent.com/dwpi/updates/ntis\\_us.html](http://www.derwent.com/dwpi/updates/ntis_us.html).) (Updated on 25-  
 MAR-2003 to correct pf field.)  
 XX SQ Sequence 341 AA;  
 Query Match 47.2%; Score 42; DB 2; Length 341;  
 Best Local Similarity 53.8%; Pred. No. 48;  
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 MTFRAQEGAFITG 14  
 : || : ||| |  
 DB 135 VVFRVEQGAFLAG 147  
 RESULT 14  
 AAR61758  
 ID AAR61758 standard; protein; 341 AA.  
 AC AAR61758;  
 XX AC  
 XX DT 09-SEP-1998 (first entry)  
 DE B. burgdorferi antigenic protein p39 beta.  
 XX KW Borrelia burgdorferi; antigenic protein; p39 alpha; p39 beta;  
 KW Lyme borreliosis; Lyme disease.  
 XX OS Borrelia burgdorferi.  
 XX PN US5780041-A.  
 XX PD 14-JUL-1998.  
 XX PF 01-MAR-1995; 95US-00396957.  
 XX PR 05-MAR-1990; 90US-00487716.  
 XX PR 05-MAR-1991; 91US-00664731.  
 XX PR 19-FEB-1993; 93US-00020245.  
 (USSH ) US DEPT HEALTH & HUMAN SERVICES.





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 20:43:36 ; Search time 15 Seconds  
(without alignments)  
109.017 Million cell updates/sec

Title: US-09-982-259B-7

Perfect score: 89

Sequence: 1 GMTFRAQEGAFITGAAC 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72	80.9	339	2 F70147	basic membrane pro
2	50	56.2	350	1 C70009	ABC transporter (1
3	47	52.8	591	2 D64204	membrane lipoprote
4	47	52.8	657	2 S73428	probable lipoprote
5	45	50.6	626	1 NDB561	exonuclease (EC 3.
6	43	48.3	256	2 E72574	probable membrane
7	43	48.3	525	2 C82914	conserved hypotet
8	43	48.3	593	2 S45281	coagulation factor
9	42	47.2	341	2 E70147	basic membrane lip
10	42	47.2	331	2 T40299	cystathionine beta
11	42	47.2	359	2 F72418	basic membrane pro
12	42	47.2	461	2 T02058	UMP synthase - com
13	42	47.2	1116	2 A84046	DNA polymerase III
14	41	46.1	194	2 B83007	conserved hypotet
15	41	46.1	222	1 CERCFE	cell division ATP-
16	41	46.1	222	2 H91167	cell division ATP-
17	41	46.1	222	2 H86013	probable lipoprote
18	41	46.1	357	2 D96986	aspartate transami
19	41	46.1	410	2 A40658	cytochrome P450 2K
20	41	46.1	504	1 S45644	conserved hypotet
21	40	44.9	460	2 S39715	glycerol kinase -
22	40	44.9	432	2 F70339	hypothetical membr
23	40	44.9	524	2 D82944	proline permease (
24	40	44.9	537	2 A75123	malate dehydrogena
25	40	44.9	604	2 S53351	probable ABC trans
26	39	43.8	359	2 F95406	probable benzocate
27	39	43.8	433	2 F75566	p-selectin precurs
28	39	43.8	646	2 JN0473	hypothetical prote
29	39	43.8	5138	2 B96695	

30 38.5 43.3 629 2 I64078  
31 38 42.7 313 2 E5853  
32 38 42.7 360 2 H70147  
33 38 42.7 418 2 S76669  
34 38 42.7 420 2 A71327  
35 38 42.7 662 2 E83201  
36 38 42.7 1548 2 S34583  
37 37.5 42.1 617 2 A71284  
38 37.5 42.1 617 2 A71284  
39 37.5 42.1 2458 2 T17420  
40 37 41.6 104 2 C90958  
41 37 41.6 104 2 D85806  
42 37 41.6 134 2 AD2592  
43 37 41.6 138 2 B90125  
44 37 41.6 145 2 F97374  
45 37 41.6 195 2 D81030

## ALIGNMENTS

### RESULT 1

F70147  
basic membrane protein A (bmpA) - Lyme disease spirochete  
N:Alternate names: antigen P39; membrane lipoprotein A  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C>Date: 13-Feb-1998 #sequence: revision 13-Feb-1998 #text\_change 29-Sep-1999  
C:Accession: F70147; I40289; I40241  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Whitson, D.; Peterson, J.; Kervage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.  
A:Reference number: A70100; MUID:9805943; PMID:9403685  
A:Accession: F70147  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-339 <KLE>  
A:Cross-references: GB:AE001143; GB:AE000783; NID:G2688279; PIDN:AAC66757.1; PID:G26882  
A:Experimental source: strain H31  
R:Simpson, W.J.; Cieplak, W.  
FEMS Microbiol. Lett. 119, 381-388, 1994  
A:Title: Nucleotide sequence and analysis of the gene in Borrelia burgdorferi encoding  
A:Reference number: I40289; MUID:94327086; PMID:8050720  
A:Accession: I40289  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-339 <RES>  
A:Cross-references: GB:L24194; NID:G508420; PIDN:AAA72406.1; PID:G508421  
R:Ojaimi, C.; Davidson, B.E.; Saint Girons, I.; Old, I.G.  
Microbiology 140, 2931-2940, 1994  
A:Title: Conservation of gene arrangement and an unusual organization of rRNA genes in  
A:Reference number: I40241; MUID:9511614; PMID:7812434  
A:Accession: I40241  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 248-339 <RE2>  
A:Cross-references: GB:L35050; NID:G516591; PIDN:AAC41401.1; PID:G516592  
C:Genetics:  
A:Gene: bmpA  
C:Superfamily: basic membrane protein C

Query Match 80.9%; Score 72; DB 2; Length 339;  
Best Local Similarity 100.0%; Pred. No. 2.6e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMTFRAQEGAFITG 14

Db 129 GMTFRAQEGAFITG 142

### RESULT 2

## C73009

ABC transporter (lipoprotein) homolog yufN - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
 C:Accession: C73009  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertoni, A.; Ehrlich, S.D.; Brouillet, S.; Bruschini, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chao, N.; Chao, N.; Emerson, P.T.; Entian, K.D.; Ewing, K.D.; Fabbet, C.; Ferrari, E.; Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Leuber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue, Y.; M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel, R.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron, T.; Winters, P.; Wipat, A.; Yamamoto, K.; Yamane, K.; Yaguchi, V.; Uchiyama, M.; Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Yata, K.; Yoshida, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A69580; MUID:98044033; PMID:9384377  
 A:Accession: C73009  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-350 <KUN>  
 A:Cross-references: GB:Z99120; GB:AL009126; NID:G2635613; PIDN:CABIS143.1; PID:G2635650  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: yufN  
 C:Superfamily: ABC transporter yufN

Query Match 56.2%; Score 50; DB 1; Length 350;  
 Best Local Similarity 60.0%; Pred. No. 0.3;  
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 MTFRAQEGAFITGAA 16  
 :||:|||||  
 Db 134 ITFKQEGSFLTGVA 148

## RESULT 3

membrane lipoprotein tmcP homolog - Mycoplasma genitalium  
 C:Species: Mycoplasma genitalium  
 C:Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 15-Sep-2000  
 C:Accession: D64204  
 R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.; M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.; C.A.; Venter, J.C.  
 Science 270, 397-403, 1995  
 A:Title: The minimal gene complement of Mycoplasma genitalium.  
 A:Reference number: A64200; MUID:96026346; PMID:7565993  
 A:Accession: D64204  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-591 <TIGR>  
 A:Cross-references: GB:J39693; NID:I43967; NID:G3844642; PIDN:AACT71256.1; PID:G1045712; T  
 A:Experimental source: strain G-37  
 C:Genetics:  
 A:Gene: tmcP  
 C:Superfamily: Mycoplasma pneumoniae probable lipoprotein D09\_orf657

Query Match 52.8%; Score 47; DB 2; Length 591;  
 Best Local Similarity 53.3%; Pred. No. 1.8;  
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 MTFRAQEGAFITGAA 16  
 :||:|||||  
 Db 180 VAFRSDGSLTGVA 194

## RESULT 4

S73428  
 C:Species: Aeropyrum pernix

probable lipoprotein D09\_orf657 - Mycoplasma pneumoniae (strain ATCC 29342)  
 A:Alternate names: MG040 homolog D09\_orf657  
 C:Species: Mycoplasma pneumoniae  
 A:Variety: ATCC 29342  
 C:Date: 26-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 15-Sep-2000  
 C:Accession: S73428  
 R:Hammelreich, R.; Hilbert, H.; Plagens, H.; Pirki, E.; Li, B.C.; Herrmann, R.  
 Nucleic Acids Res. 24, 4420-4449, 1996  
 A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
 A:Reference number: S73327; MUID:97105885; PMID:8948633  
 A:Accession: S73428  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-657 <HIM>  
 A:Cross-references: EMBL:AB000012; GB:U00089; NID:G1673755; PIDN:AA095750.1; PID:G16737  
 A:Note: The nucleotide sequence was submitted to the EMBL Data Library, November 1996  
 C:Genetics:  
 A:Genetic code: SGC3  
 C:Superfamily: Mycoplasma pneumoniae probable lipoprotein D09\_orf657

Query Match 52.8%; Score 47; DB 2; Length 657;  
 Best Local Similarity 53.3%; Pred. No. 2;  
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 MTFRAQEGAFITGAA 16  
 :||:|||||  
 Db 176 VAFRSDGSLTGVA 190

## RESULT 5

NDBE61  
 exonuclease (EC 3.1.11.1) - human herpesvirus 1 (strain 17)  
 A:Alternate names: gene UL12 protein (deoxyribonuclease)  
 C:Species: human herpesvirus 1  
 A:Note: host Homo sapiens (man)  
 C:Date: 04-Dec-1986 #sequence\_revision 04-Dec-1986 #text\_change 16-Jun-2000  
 C:Accession: A00781; C30083  
 R:McGeoch, D.J.; Dolan, A.; Frame, M.C.  
 Nucleic Acids Res. 14, 3435-3448, 1986  
 A:Title: DNA sequence of the region in the genome of herpes simplex virus type 1 contain  
 A:Reference number: A93620; MUID:86205244; PMID:3010237  
 A:Accession: A00781  
 A:Molecule type: DNA  
 A:Residues: 1-626 <MCQ>  
 A:Cross-references: GB:X03839; NID:G59841; PIDN:CAA27453.1; PID:G59844  
 R:McGeoch, D.J.; Dairymple, M.A.; Davison, A.J.; Dolan, A.; Frame, M.C.; McNab, D.; Per  
 J. Gen. Virol. 69, 1531-1574, 1988  
 A:Title: The complete DNA sequence of the long unique region in the genome of herpes si  
 A:Reference number: A30083; MUID:88274327; PMID:2839594  
 A:Accession: C30083  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-626 <MC2>  
 A:Cross-references: GB:X14112; NID:G1944536; PIDN:CAA32325.1; PID:G59513; GB:D00317  
 C:Genetics:  
 A:Gene: UL12  
 A:Map position: 0.16-0.20  
 C:Superfamily: herpesvirus exonuclease  
 C:Keywords: exonuclease; hydrolase

Query Match 50.6%; Score 45; DB 1; Length 626;  
 Best Local Similarity 56.2%; Pred. No. 4.5;  
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFITGAA 16  
 :||:|||||  
 Db 537 GVTFRLEDGAGALGAA 552

## RESULT 6

E72574  
 probable membrane protein APE1877 - Aeropyrum pernix (strain K1)  
 C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 08-Sep-2000  
 C/Accession: E72574  
 R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; DNA Res. 6, 83-101, 1999  
 A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix  
 A/Reference number: A72450; MUID:99310339; PMID:10382966  
 A/Accession: E72574  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-256 <KAW>  
 A/Cross-references: DDBJ:AP000062; NID:G5105244; PIDN:BA080882.1; PID:G5105569  
 A/Experimental source: strain K1  
 C/Genetics:  
 A/Gene: APE1877  
 C/Superfamily: Streptomyces coelicolor probable integral membrane protein SC6G10.12

Query Match 48.3%; Score 43; DB 2; Length 256;  
 Best Local Similarity 56.2%; Pred. No. 4.1;  
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 GMTFRAOEGAFITGAA 16  
 :|||:|||||  
 Db 217 GIAPFAHVGFLTGVA 232

RESULT 7  
 C82914  
 conserved hypothetical membrane lipoprotein U0226 [imported] - Ureaplasma urealyticum  
 C/Species: Ureaplasma urealyticum  
 C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
 C/Accession: C82914  
 R/Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H. submitted to GenBank, February 2000  
 A/Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mitochondrial genome  
 A/Reference number: A82870  
 A/Accession: C82914  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-525 <GLA>  
 A/Cross-references: GB:AF002123; GB:AF222894; NID:G6899229; PIDN:AAF30635.1; GSPDB:GN001  
 A/Experimental source: serovar 3; biovar 1  
 C/Genetics:  
 A/Gene: U0226  
 A/Genetic code: SGC3

Query Match 48.3%; Score 43; DB 2; Length 525;  
 Best Local Similarity 53.3%; Pred. No. 8.7;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 MTFRAGGAFITGAA 16  
 :|||:|||||  
 Db 128 ITRYADQAFITGAA 142

RESULT 8  
 S45281  
 coagulation factor XIIa (EC 3.4.21.38) precursor - bovine (fragment)  
 N/Alternate names: Hageman factor (activated)  
 C/Species: Bos primigenius taurus (cattle)  
 C/Date: 10-Apr-1995 #sequence\_revision 22-Apr-1995 #text\_change 21-Jan-2000  
 C/Accession: S45281; A61329  
 R/Shibuya, Y.; Semba, U.; Okabe, H.; Kambara, T.; Yamamoto, T. Biochim. Biophys. Acta 1206, 63-70, 1994  
 A/Title: Primary structure of bovine Hageman factor (blood coagulation factor XII): complete sequence  
 A/Reference number: S45281; MUID:94242782; PMID:8186251  
 A/Accession: S45281  
 A/Molecule type: mRNA  
 A/Residues: 1-593 <SHI>  
 A/Cross-references: GB:S70164  
 A/Note: the authors translated the codon GAG for residue 23 as Val, GAG for residue 70 as Pro, CTC for residue 203 as Phe, GTG for residue 247 as Leu, CCG for residue 286 as Gly, and ATC for residue 505 as Leu

R:Fujikawa, K.; Walsh, K.A.; Davie, E.W. Biochemistry 16, 2270-2278, 1977  
 A/Title: Isolation and characterization of bovine factor XII (Hageman factor).  
 A/Reference number: A61329; MUID:77182112; PMID:861210  
 A/Accession: A61329  
 A/Molecule type: protein  
 A/Residues: 10-16, 'X', 18-19; 525-550 <FUI>  
 C/Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology  
 C/Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; monomer; plasma;  
 F:37-78/Domain: fibronectin type II repeat homology <IF2>  
 F:88-120/Domain: EGF homology <EGF>  
 F:125-160/Domain: fibronectin type I repeat homology <FBI>  
 F:207-287/Domain: kringle homology <KRG>  
 F:350-587/Domain: trypsin homology <TRY>  
 F:541/Active site: Ser #status predicted

Query Match 48.3%; Score 43; DB 2; Length 593;  
 Best Local Similarity 41.2%; Pred. No. 9.9;  
 Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 GMTFRAOEGAFITGAA 17  
 :|||:|||||  
 Db 214 GLSYRGAGTILSGAPC 230

RESULT 9  
 E70147  
 basic membrane lipoprotein B (bnpB) - Lyme disease spirochete  
 C/Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C/Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 29-Sep-1999  
 C/Accession: E70147; I40290; I40242  
 R/Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Whitson, D.; Peterson, J.; Keriavala, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B. Nature 390, 580-586, 1997  
 A/Authors: Smith, H.O.; Venter, J.C.  
 A/Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.  
 A/Reference number: A70100; MUID:98065943; PMID:9403685  
 A/Accession: E70147  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-341 <KLE>  
 A/Cross-references: GB:AF001143; GB:AB000783; NID:G2688279; PIDN:AAC66758.1; PID:G2688;  
 A/Experimental source: strain B31  
 R/Simpson, W.J.; Ciepiak, W. FEMS Microbiol. Lett. 119, 381-388, 1994  
 A/Title: Nucleotide sequence and analysis of the gene in Borrelia burgdorferi encoding  
 A/Reference number: I40289; MUID:94327086; PMID:8050720  
 A/Accession: I40290  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-232, 'A', 234-317, 'V', 319-341 <RES>  
 A/Cross-references: GB:I24194; NID:G508423; PIDN:AAA72407.1; PID:G508422  
 R/Ojaimi, C.; Davidson, B.E.; Saint Girons, I.; O'd, I.G. Microbiology 140, 2931-2940, 1994  
 A/Title: Conservation of gene arrangement and an unusual organization of rRNA genes in  
 A/Reference number: I40241; MUID:9511614; PMID:7812434  
 A/Accession: I40242  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-44, 'A', 46-179 <RE2>  
 A/Cross-references: GB:I35050; NID:G516591; PIDN:AAC41402.1; PID:G551744  
 C/Superfamily: basic membrane protein C

Query Match 47.2%; Score 42; DB 2; Length 341;  
 Best Local Similarity 53.8%; Pred. No. 8.5;  
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 MTFRAGGAFITG 14  
 :|||:|||||  
 Db 135 VVFRVQGAFLAG 147

## RESULT 10

T40299  
Cystathionine beta-synthase (EC 4.2.1.22) SPBC36.04 [similarity] - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 20-Jun-2000  
C:Accession: T40299  
R:Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Moestl, D.; Duesterh submitted to the EMBL Data Library, May 1998  
A:Reference number: Z11919  
A:Accession: T40299  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-351 <LYN>  
A:Cross-references: EMBL:AL023589; PIDN:CAAL19052.1; GSPDB:GN00067; SPDB:SPBC36.04  
A:Experimental source: strain 972h; cosmid c36  
C:Genetics:  
A:Gene: SPDB:SPBC36.04  
A:Map position: 2  
A:Introns: 45/1  
C:Superfamily: threonine dehydratase  
C:Keywords: carbon-oxygen lyase; hydro-lyase; phosphoprotein; pyridoxal phosphate  
P:54/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 47.2%; Score 42; DB 2; Length 351;  
Best Local Similarity 55.6%; Pred. No. 8.8;  
Matches 10; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

QY 2 MTFRA--QEGAFITGAAC 17

Db 274 MFFRLDQEGFLGSSSC 291

## RESULT 11

F72418  
basic membrane protein - Thermotoga maritima (strain MSB8)  
C:Species: Thermotoga maritima  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: F72418  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D. C.M.  
Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing  
A:Reference number: A72200; MUID:99287316; PMID:10360571  
A:Accession: F72418  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-359 <ARN>  
A:Cross-references: GB:AE001696; GB:AE000512; NID:94980582; PIDN:AAD35196.1; PID:9498059  
A:Experimental source: strain MSB8  
C:Genetics:  
A:Gene: TW0102  
C:Superfamily: ABC transporter yufN

Query Match 47.2%; Score 42; DB 2; Length 359;  
Best Local Similarity 66.7%; Pred. No. 9;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 TFRAGGAFITG 14

Db 122 TFKQGEAEFLVG 133

## RESULT 12

T02058  
UMP synthase - common tobacco (fragment)  
N:Alternate names: uridine monophosphate  
N:Contains: orotate phosphoribosyltransferase  
C:Species: Nicotiana tabacum (common tobacco)  
C>Date: 26-Feb-1999 #sequence\_revision 26-Feb-1999 #text\_change 18-Jun-1999  
C:Accession: T02058  
R:Thornburg, R.W.  
submitted to the EMBL Data Library, March 1995

A:Reference number: Z14527

A:Accession: T02058  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-461 <THO>  
A:Cross-references: EMBL:U22260; NID:g747979; PIDN:AAC49115.1; PID:g747980  
A:Experimental source: tissue-type leaf  
C:Genetics:  
A:Gene: pyrS-6  
C:Function:  
A:Description: catalyzes conversion of orotidine 5'-phosphate and diphosphate to orotate and CO<sub>2</sub>  
A:Pathway: pyrimidine metabolism  
C:Superfamily: UMP synthase; orotate phosphoribosyltransferase homology; orotidine-5'-P  
C:Keywords: carbon-carbon lyase; carboxy-lyase; glycosyltransferase; pentosyltransferase

Query Match 47.2%; Score 42; DB 2; Length 461;  
Best Local Similarity 60.0%; Pred. No. 12;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 TFRAGGAFITGAAC 17

Db 88 TAKAIEGAFKPGQAC 102

## RESULT 13

A84046  
DNA polymerase III alpha subunit dnaB [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: A84046  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: A84046  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1116 <STO>  
A:Cross-references: GB:AP001517; GB:BA000004; NID:G10175500; PIDN:BA06888.1; GSPDB:GN0  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: dnaB  
C:Superfamily: DNA-directed DNA polymerase III alpha chain

Query Match 47.2%; Score 42; DB 2; Length 1116;  
Best Local Similarity 60.0%; Pred. No. 29;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFITGA 15

Db 685 GEALRQGEAEFVTGA 699

## RESULT 14

B83007  
conserved hypothetical protein PA5115 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: B83007  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim .; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path  
A:Reference number: AB2950; MUID:20437337; PMID:10984043  
A:Accession: B83007  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-194 <STO>  
A:Cross-references: GB:AE004924; GB:AE004091; NID:g9951407; PIDN:AAG08500.1; GSPDB:GN00  
A:Experimental source: strain PA01  
C:Genetics:

A:Gene: PA5115

Query Match 46.1%; Score 41; DB 2; Length 194;  
 Best Local Similarity 80.0%; Pred. No. 7.2;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 EGAFLTGAAC 17  
 ||| |||||  
 Db 176 EGRFHTGAAC 185

## RESULT 15

CEECFE  
 cell division ATP-binding protein ftsE - Escherichia coli (strain K-12)  
 C:Species: Escherichia coli  
 C:Date: 31-Dec-1990 #sequence revision 31-Dec-1990 #text\_change 01-Mar-2002  
 C:Accession: S01131; S47682; B65143  
 R:Gill, D.R.; Hatfull, G.F.; Salmond, G.P.C.  
 Mol. Gen. Genet. 205, 134-145, 1986  
 A:Title: A new cell division operon in Escherichia coli.  
 A:Reference number: S03129; MUID:87089083; PMID:3025556  
 A:Accession: S01131  
 A:Molecule type: DNA  
 A:Residues: 1-222 <GIL>  
 A:Cross-references: EMBL:X04398; NID:G41496; PIDN:CAA27985.1; PID:G41499  
 R:Plunkett, G.  
 submitted to the EMBL Data Library, March 1994  
 A:Reference number: S47666  
 A:Accession: S47682  
 A:Molecule type: DNA  
 A:Residues: 1-222 <PLJ>  
 A:Cross-references: EMBL:U00039; NID:G466582; PIDN:AAB18438.1; PID:G466599  
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co-  
 .A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A:Title: The complete genome sequence of Escherichia coli K-12.  
 A:Reference number: A64720; MUID:97426617; PMID:9278503  
 A:Accession: B65143  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-222 <BLAT>  
 A:Cross-references: GB:AF000422; GB:U00036; NID:G1789868; PIDN:AAC76488.1; PID:G1789873;  
 A:Experimental source: strain K-12, substrain MG1655  
 C:Genetics:  
 A:Gene: ftsE  
 A:Map position: 76 min  
 C:Superfamily: short-chain ATP-binding cassette proteins; ATP-binding cassette homology  
 C:Keywords: ATP; cell division; inner membrane; nucleotide binding; P-loop  
 F:18-213/Domain: ATP-binding cassette homology <ABC>  
 F:35-42/Region: nucleotide-binding motif A (P-loop)  
 F:158-163/Region: nucleotide-binding motif B

Query Match 46.1%; Score 41; DB 1; Length 222;  
 Best Local Similarity 62.5%; Pred. No. 8.3;  
 Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 1 GMTFRAQEG--AFLTG 14  
 :||| |||||  
 Db 20 GVTFFMQPGEMAFLTG 35

Search completed: October 6, 2004, 20:49:25  
 Job time : 18 secs

GenCore version: 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 20:40:45 ; Search time 9 Seconds  
(without alignments)  
98.355 Million cell updates/sec

Title: US-09-982-259B-7

Perfect score: 89

Sequence: 1 GMTFRAQBGAFITGAAC 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72	80.9	325	1	BMPA_BORGA
2	72	80.9	339	1	BMPA_BORAF
3	72	80.9	339	1	BMPA_BORBU
4	50	56.2	350	1	YUFN_BACSU
5	47	52.8	341	1	BMPB_BORAF
6	47	52.8	341	1	BMPB_BORGA
7	47	52.8	591	1	Y040_MYCGE
8	47	52.8	657	1	Y040_MYCPN
9	45	50.6	626	1	EXON_HSV11
10	43	48.3	593	1	FAL12_BOVIN
11	42	47.2	341	1	BMPB_BORBU
12	42	47.2	448	1	Y4QG_RHISN
13	42	47.2	461	1	PVR5_TORAC
14	42	47.2	1116	1	DP3A_BACHD
15	41	46.1	222	1	FTSE_ECOLI
16	41	46.1	410	1	AAB1_RHIME
17	41	46.1	410	1	AAB2_RHIME
18	41	46.1	491	1	CPK3_ONCMY
19	41	46.1	504	1	CPK1_ONCMY
20	40	44.9	460	1	YWDJ_BACSU
21	40	44.9	492	1	GLPK_AQUAE
22	40	44.9	604	1	MAON_HUMAN
23	39	43.8	337	1	RIR2_TRYBB
24	39	43.8	338	1	RIR2_DICDI
25	39	43.8	379	1	DNAJ_PASHA
26	39	43.8	646	1	LEM3_BOVIN
27	39	43.8	769	1	LEM3_SHEEP
28	38.5	43.3	629	1	GIDA_HAZIN
29	38	42.7	326	1	VS09_ROTHX
30	38	42.7	341	1	BMPD_BORBU
31	38	42.7	345	1	RIR2_LEIAM
32	38	42.7	377	1	DNAJ_HABDU
33	38	42.7	420	1	SECF_TREPA

34	38	42.7	1877	1	PKK5_MOUSE
35	37	41.6	392	1	YFCJ_ECOLI
36	37	41.6	392	1	YFCJ_ECOLI
37	37	41.6	392	1	YFCJ_SHIFL
38	37	41.6	485	1	LEM2_BOVIN
39	37	41.6	576	1	PGPL_HUMAN
40	37	41.6	600	1	SYR_RALSO
41	37	41.6	647	1	ACSA_XANAC
42	37	41.6	647	1	ACSA_XANCP
43	36.5	41.0	496	1	GLX2_THEMA
44	36.5	41.0	499	1	MEP2_YEAST
45	36	40.4	178	1	CPT_STEVL

## ALIGNMENTS

### RESULT 1

ID	BMPA_BORGA	STANDARD;	PRT;	325 AA.
AC	O31357; O31360;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Basic membrane protein A precursor (Immunodominant antigen P39)			
DE	(Fragment).			
GN	BMPA.			
OS	Borrelia garinii.			
OC	Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI TaxID=29519;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=PLI, and PLI;			
RX	MEDLINE=98010210; PubMed=9350727;			
RA	Roesler D., Hauser U., Wilske B.;			
RT	"Heterogeneity of BmpA (P39) among European isolates of Borrelia burgdorferi sensu lato and influence of interspecies variability on serodiagnosis."			
RT	Clin. Microbiol. 35:2752-2758(1997).			
CC	FUNCTION: NOT KNOWN; IMMUNOGENIC PROTEIN.			
CC	SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).			
CC	SIMILARITY: Belongs to the BMP lipoprotein family.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL; X97244; CAA65883.1; -			
CC	EMBL; X97238; CAA65877.1; -			
DR	InterPro; IPR003760; Bmp.			
DR	InterPro; IPR000437; Prok_lipoprot_S.			
DR	Pfam; PF02608; Bmp; 1			
DR	PROSITE; PS00013; PROKAR_LIPOPROTEIN; PARTIAL.			
KW	Antigen; Membrane; Lipoprotein; Signal; Palmitate.			
FT	NON TER	1	1	
FT	SIGNAL	<1	3	PROBABLE.
FT	CHAIN	4	325	BASIC MEMBRANE PROTEIN A.
FT	LIPID	4	4	N-palmitoyl cysteine (Probable).
FT	LIPID	4	4	S-diacylglycerol cysteine (Probable).
FT	VARIANT	8	8	G -> D (IN STRAIN PLI).
FT	VARIANT	12	12	S -> N (IN STRAIN PLI).
FT	VARIANT	93	93	P -> S (IN STRAIN PLI).
FT	VARIANT	111	111	T -> A (IN STRAIN PLI).
FT	VARIANT	180	180	N -> D (IN STRAIN PLI).
FT	VARIANT	205	205	S -> G (IN STRAIN PLI).
FT	VARIANT	239	239	I -> V (IN STRAIN PLI).
FT	VARIANT	243	243	A -> S (IN STRAIN PLI).
FT	VARIANT	250	250	S -> A (IN STRAIN PLI).

FT VARIANT 257 257 N -> S (IN STRAIN PL1).  
SQ SEQUENCE 325 AA; 35480 MW; D19281A2A4B1C158 CRC64;

Query Match 80.9%; Score 72; DB 1; Length 325;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFITG 14  
|||||  
DB 115 GMTFRAQEGAFITG 128

## RESULT 2

BMPA BORAF  
ID BMPA\_BORAF STANDARD; PRT; 339 AA.  
AC Q31280; O31281; O31282; O31283;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Basic membrane protein A precursor (Immunodominant antigen P39).  
GN BMPA.  
OS Borrelia afzelii.  
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=29518;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PKO, PWU1, PLE, and PLJ7;  
RX MEDLINE=98010210; PubMed=9350727;  
RA Roessler D., Hauser U., Wilske B.;  
RT "Heterogeneity of BmpA (P39) among European isolates of Borrelia burgdorferi sensu lato and influence of interspecies variability on serodiagnosis.";  
RL J. Clin. Microbiol. 35:2752-2758(1997).  
CC -!- FUNCTION: NOT KNOWN; IMMUNOGENIC PROTEIN.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).  
CC -!- SIMILARITY: Belongs to the BMP lipoprotein family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X81516; CAA57236.1; -;  
EMBL; X97237; CAA65876.1; -;  
EMBL; X97239; CAA65878.1; -;  
EMBL; X97241; CAA65880.1; -;  
InterPro: IPR003763; Bmp.  
InterPro: IPR000437; Prok\_lipoprot\_S.  
Pfam: PF02608; Bmp; 1.  
PROSITE: PS00013; PROKAR\_LIPOPROTEIN; FALSE NEG.  
Antigen; Membrane; Lipoprotein; Signal; Palmitate.  
FT SIGNAL 1 17 PROBABLE.  
FT CHAIN 18 339 BASIC MEMBRANE PROTEIN A.  
FT LIPID 18 18 N-palmitoyl cysteine (Probable).  
FT LIPID 18 18 S-diacylglycerol cysteine (Probable).  
FT VARIANT 125 125 A -> S (IN STRAIN PL1).  
FT VARIANT 214 214 I -> T (IN STRAIN PLJ7).  
FT VARIANT 229 229 A -> P (IN STRAIN PKO).  
FT VARIANT 254 254 I -> V (IN STRAIN PL7).  
FT VARIANT 268 268 L -> F (IN STRAINS PLE AND PKO).  
SQ SEQUENCE 339 AA; 36966 MW; 1BEABDF8CA06FB4 CRC64;

Query Match 80.9%; Score 72; DB 1; Length 339;  
Best Local Similarity 100.0%; Pred. No. 2e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFITG 14  
|||||  
DB 129 GMTFRAQEGAFITG 142

## RESULT 3

BMPA BORBU  
ID BMPA\_BORBU STANDARD; PRT; 339 AA.  
AC Q45010; P94249; Q44857;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Basic membrane protein A precursor (Immunodominant antigen P39).  
GN BMPA OR BH0383.  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sh-2-82;  
RX MEDLINE=94327086; PubMed=8050720;  
RA Simpson W.J., Cieplak W., Schrumpf M.E., Barbour A.G., Schwan T.G.;  
RT "Nucleotide sequence and analysis of the gene in Borrelia burgdorferi encoding the immunogenic P39 antigen.";  
RL FEMS Microbiol. Lett. 119:381-388(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=297;  
RX MEDLINE=97132632; PubMed=8978084;  
RA Aron L., Toth C., Godfrey H.P., Cabello F.C.;  
RT "Identification and mapping of a chromosomal gene cluster of Borrelia burgdorferi containing genes expressed in vivo.";  
RL FEMS Microbiol. Lett. 145:309-314(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BL206;  
RA Orlova T., Burgysheva J., Novikova S., Godfrey H.P., Cabello F.C.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 35210 / B31;  
RX MEDLINE=98065943; PubMed=9403685;  
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A., Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M., Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D., Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M., van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J., Uitterback T., Matthey L., McDonald L., Artiach P., Bowman C., Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B., Smith H.O., Venter J.C.;  
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.";  
RL Nature 390:580-586(1997).  
RN [5]  
RP SEQUENCE OF 248-339 FROM N.A.  
RC STRAIN=212;  
RX MEDLINE=95111614; PubMed=7812434;  
RA Ojalimi C., Davidson B.E., Saint-Girons I., Old I.G.;  
RT "Conservation of gene arrangement and an unusual organization of rRNA genes in the linear chromosomes of the Lyme disease spirochaetes Borrelia burgdorferi, B. garinii and B. afzelii.";  
RL Microbiology 140:2931-2940(1994).  
CC -!- FUNCTION: NOT KNOWN; IMMUNOGENIC PROTEIN.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).  
CC -!- SIMILARITY: Belongs to the BMP lipoprotein family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; L24194; AAA72406.1; -;

DR EMBL; U49938; AAC44712.1; --  
 DR EMBL; AF400111; AAM89914.1; --  
 DR EMBL; AE001143; AAC66757.1; --  
 DR EMBL; L35050; AAC41401.1; --  
 DR PIR; F70147; F70147.  
 DR TIGR; BB0383; --  
 DR InterPro; IPR003760; Bmp.  
 DR InterPro; IPR000437; Prok\_lipoprot\_s.  
 DR PROSITE; PS00013; PROKAR\_LIPOPROTEIN; FALSE NEG.  
 KW Antigen; Membrane; Lipoprotein; Signal; Complete proteome; Palmitate.  
 FT SIGNAL 1 17 PROBABLE.  
 FT CHAIN 18 339 BASIC MEMBRANE PROTEIN A.  
 FT LIPID 18 18 N-palmitoyl cysteine (Probable).  
 FT LIPID 18 18 S-diacylglycerol cysteine (Probable).  
 FT VARIANT 278 278 F -> S (IN STRAIN 297).  
 SQ SEQUENCE 339 AA; 36968 MW; A3581868CA7DB923 CRC64;  
 Query Match 80.9%; Score 72; DB 1; Length 339;  
 Best Local Similarity 100.0%; Pred. No. 2e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GMTFRAQEGAFLTG 14  
 DB 129 GMTFRAQEGAFLTG 142  
 RESULT 4  
 YUEN\_BACSU STANDARD; PRT; 350 AA.  
 AC 005252;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Hypothetical lipoprotein yuFN precursor.  
 GN YUEN OR BSU31540.  
 CS Bacillus subtilis.  
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=97419515; PubMed=9274030;  
 RA Oudega B., Koningsstein G., Rodrigues L., de Sales Ramon M.,  
 RA Hilbert H., Duesterhoeft A., Pohl T.M., Weitzenecker T.;  
 RA "Analysis of the Bacillus subtilis genome: cloning and nucleotide  
 RT sequence of a 62 kb region between 275 degrees (rmb) and 284 degrees  
 RT (pai)";  
 RL Microbiology 143:2769-2774(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst E., Ogasawara N., Mszerski I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
 RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 RA Ghm S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
 RA Giuseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue V.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Ncback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
 RA Presecan E., Pujic R., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,

RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
 RA Toseato V., Uchiyama S., Vandenberg M., Vannier F., Vassarotti A.,  
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamahe K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,  
 RT "The complete sequence of the Gram-positive bacterium Bacillus  
 RT subtilis";  
 RL Nature 390:249-256(1997).  
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor  
 CC (Potential).  
 CC -1- SIMILARITY: Belongs to the BMP lipoprotein family.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; Z3937; CAB07936.1; --  
 CC EMBL; Z99120; CAB15143.1; --  
 CC PIR; C70009; C70009.  
 CC Subtilist; BG12349; yuFN.  
 CC InterPro; IPR003760; Bmp.  
 CC InterPro; IPR000437; Prok\_lipoprot\_s.  
 CC Pfam; PF02608; Bmp; 1.  
 CC PROSITE; PS00013; PROKAR\_LIPOPROTEIN; FALSE NEG.  
 KW Hypothetical protein; Membrane; Lipoprotein; Signal;  
 KW Complete proteome; Palmitate.  
 FT SIGNAL 1 13 PROBABLE.  
 FT CHAIN 14 350 HYPOTHETICAL LIPOPROTEIN YUFN.  
 FT LIPID 14 14 N-palmitoyl cysteine (Probable).  
 FT LIPID 14 14 S-diacylglycerol cysteine (Probable).  
 SQ SEQUENCE 350 AA; 37349 MW; 16D5176A52A99284 CRC64;  
 Query Match 56.2%; Score 50; DB 1; Length 350;  
 Best Local Similarity 60.0%; Pred. No. 0.19;  
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 MTFRAQEGAFLTGAA 16  
 DB 134 MTFRAQEGAFLTGVA 148  
 RESULT 5  
 BMPB BORAF STANDARD; PRT; 341 AA.  
 ID BMPB BORAF  
 AC Q31284;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Basic membrane protein B precursor.  
 GN BMPB.  
 OS Borrelia afzelii.  
 CC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=29518;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PKO;  
 RX MEDLINE=98010210; PubMed=9350727;  
 RA Roessler D., Hauser U., Wilske B.;  
 RT "Heterogeneity of BmpA (p39) among European isolates of Borrelia  
 RT burgdorferi sensu lato and influence of interspecies variability on  
 RT serodiagnosis";  
 RL J. Clin. Microbiol. 35:2752-2758(1997).  
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor  
 CC (Probable).  
 CC -1- SIMILARITY: Belongs to the BMP lipoprotein family.



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X81519; CAA57239.1; --  
 DR InterPro; IPR003760; Bmp.  
 DR InterPro; IPR000437; Prok\_lipoprot\_S.  
 DR Pfam; PF02608; Bmp; 1.  
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.  
 KW Membrane; Lipoprotein; Signal; Palmitate.  
 FT SIGNAL 1 14 PROBABLE.  
 FT CHAIN 15 341 BASIC MEMBRANE PROTEIN B.  
 FT LIPID 15 15 N-palmitoyl cysteine (Probable).  
 FT LIPID 15 15 S-diacylglycerol cysteine (Probable).  
 SQ SEQUENCE 341 AA; 37198 MW; 1E8EA3E3D54ACDB1 CRC64;

Query Match 52.8%; Score 47; DB 1; Length 341;  
 Best Local Similarity 57.1%; Pred. No. 0.65;  
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFITG 14  
 | : : : : : |  
 Db 134 GVVFRIEQGAFITG 147

## RESULT 6

ID BMBP\_BORGA STANDARD; PRT; 341 AA.  
 AC Q31362;  
 DT 15-DEC-1998 (Rel. 37, Created);  
 DT 15-DEC-1998 (Rel. 37, Last sequence update);  
 DT 10-OCT-2003 (Rel. 42, Last annotation update);  
 DE Basic membrane protein B precursor.  
 GN BMBP  
 OS Borrelia garinii.  
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=29519;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PB1.  
 RX MEDLINE=98010210; PubMed=9350727;  
 RA Roessler D., Hauser U., Wilske B.;  
 RT "Heterogeneity of BmpA (P39) among European isolates of Borrelia  
 RT burgdorferi sensu lato and influence of interspecies variability on  
 RT serodiagnosis.";  
 RL J. Clin. Microbiol. 35:2752-2758(1997).  
 CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor  
 CC (Probable).  
 CC -!- SIMILARITY: Belongs to the BMP lipoprotein family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X81518; CAA57238.1; --  
 DR InterPro; IPR003760; Bmp.  
 DR InterPro; IPR000437; Prok\_lipoprot\_S.  
 DR Pfam; PF02608; Bmp; 1.  
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.  
 KW Membrane; Lipoprotein; Signal; Palmitate.  
 FT SIGNAL 1 14 PROBABLE.  
 FT CHAIN 15 341 BASIC MEMBRANE PROTEIN B.  
 FT LIPID 15 15 N-palmitoyl cysteine (Probable).  
 FT LIPID 15 15 S-diacylglycerol cysteine (Probable).  
 SQ SEQUENCE 341 AA; 37236 MW; 442BEF0BFEDD0C3A CRC64;

Query Match 52.8%; Score 47; DB 1; Length 341;  
 Best Local Similarity 57.1%; Pred. No. 0.65;  
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFITG 14  
 | : : : : : |  
 Db 134 GVVFRIEQGAFITG 147

## RESULT 7

ID Y040 MYCGE STANDARD; PRT; 591 AA.  
 AC P47286;  
 DT 01-OCT-1996 (Rel. 34, Created);  
 DT 01-OCT-1996 (Rel. 34, Last sequence update);  
 DT 10-OCT-2003 (Rel. 42, Last annotation update);  
 DE Hypothetical lipoprotein MG040 precursor.  
 GN MG040.  
 OS Mycoplasma genitalium.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33530 / G-37;  
 RX MEDLINE=96026346; PubMed=7569993;  
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,  
 RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
 RA Tomb J.-P., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,  
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
 RT "The minimal gene complement of Mycoplasma genitalium.";  
 RL Science 270:397-403(1995).  
 RN [2]  
 RP SEQUENCE OF 448-517 FROM N.A.  
 RC STRAIN=ATCC 33530 / G-37;  
 RX MEDLINE=94075230; PubMed=8253680;  
 RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;  
 RT "A survey of the Mycoplasma genitalium genome by using random  
 RT sequencing.";  
 RL J. Bacteriol. 175:7918-7930(1993).  
 CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor  
 CC (Potential).  
 CC -!- SIMILARITY: SOME, TO T.PALLIDIUM TMPC.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U39683; AAC71256.1; --  
 DR EMBL; U02125; AAD12400.1; --  
 DR PIR; D64204; D64204.  
 DR TIGR; MG040; --  
 DR InterPro; IPR003760; Bmp.  
 DR InterPro; IPR000437; Prok\_lipoprot\_S.  
 DR Pfam; PF02608; Bmp; 1.  
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.  
 KW Hypothetical protein; Lipoprotein; Membrane; Signal;  
 KW Complete proteome; Palmitate.  
 FT SIGNAL 1 30 POTENTIAL.  
 FT CHAIN 31 591 HYPOTHETICAL LIPOPROTEIN MG040.  
 FT LIPID 31 31 N-palmitoyl cysteine (Potential).  
 FT LIPID 31 31 S-diacylglycerol cysteine (Potential).  
 SQ SEQUENCE 591 AA; 64019 MW; 14A8FA31BB7E0928 CRC64;

Query Match 52.8%; Score 47; DB 1; Length 591;  
 Best Local Similarity 53.3%; Pred. No. 1.1;  
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

```

OY      2 MTFRAQEGAFITGAA 16
      :|:|:|:|:|
DB      180 VAFRSDSGSFLTGVA 194

RESULT 8
Y040_MYCPN
ID_Y040_MYCPN      STANDARD;      PRT;      657 AA.
AC      P75062;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Hypothetical lipoprotein MG040 homolog precursor (D09_orf657).
GN      MPN052 OR MP102.
OS      Mycoplasma pneumoniae.
OC      Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX      NCBI_TaxID=2104;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 29342 / M129;
RX      MEDLINE=97105885; PubMed=8948633;
RA      Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA      Herrmann R.;
RT      "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT      pneumoniae.";
RL      Nucleic Acids Res. 24:4420-4449(1996).
CC      -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
      (Potential).
CC      -!- SIMILARITY: SOME, TO T.PALLIDIUM TMPC.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AE000012; AAB395750.1; -
PIR; S73428; S73428.
InterPro; IPR003760; Bmp.
ProSite; PS02608; Bmp; 1.
PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
KW      Hypothetical protein; Lipoprotein; Membrane; Signal;
KW      Complete proteome; Palmitate.
FT      SIGNAL.
FT      CHAIN.
FT      LIPID.
FT      LIPID.
FT      LIPID.
SQ      SEQUENCE 657 AA; 71671 MW; 6147BJA5673606A9 CRC64;

Query Match      52.8%; Score 47; DB 1; Length 657;
Best Local Similarity 53.3%; Pred. No. 1.3;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY      2 MTFRAQEGAFITGAA 16
      :|:|:|:|:|
DB      176 VAFRSDSGSFLTGVA 190

RESULT 9
EXON_HSV11
ID_EXON_HSV11      STANDARD;      PRT;      626 AA.
AC      P04234;
DT      20-MAR-1987 (Rel. 04, Created)
DT      20-MAR-1987 (Rel. 04, Last sequence update)
DT      01-MAY-1992 (Rel. 22, Last annotation update)
DE      Alkaline exonuclease (EC 3.1.11.-).
GN      UL12.
OS      Herpes simplex virus (type 1 / strain 17).
OC      Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

```

---

```

OC      Alphaherpesvirinae; Simplexvirus.
OX      NCBI_TaxID=10299;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=88274327; PubMed=2839594;
RA      McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
RA      McNab D., Perry L.J., Scott J.E., Taylor P.;
RT      "The complete DNA sequence of the long unique region in the genome of
RT      herpes simplex virus type 1.";
RL      J. Gen. Virol. 69:1531-1574 (1988).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=86205244; PubMed=3010237;
RA      McGeoch D.J., Dolan A., Frame M.C.;
RT      "DNA sequence of the region in the genome of herpes simplex virus
RT      type 1 containing the exonuclease gene and neighbouring genes.";
RL      Nucleic Acids Res. 14:3435-3448(1986).
CC      -!- SIMILARITY: Belongs to the herpesviruses alkaline exonuclease
      family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X14112; CAA32325.1; -
EMBL; X03839; CAA27453.1; -
PIR; A00781; NDBE61.
InterPro; IPR001616; Herpes alk exo.
PFam; PF01771; Herpes alk exo; 1.
PRINTS; PR00924; ALKXNUCLASE.
KW      Hydrolyase; Nuclease; Exonuclease.
SQ      SEQUENCE 626 AA; 67508 MW; 7B86C941A0105035 CRC64;

Query Match      50.6%; Score 45; DB 1; Length 626;
Best Local Similarity 56.2%; Pred. No. 2.8;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY      1 GMTFRAQEGAFITGAA 16
      :|:|:|:|:|
DB      537 GVTFRLEDGAGALGAA 552

RESULT 10
FA12_BOVIN
ID_FA12_BOVIN      STANDARD;      PRT;      593 AA.
AC      P98140;
DT      01-FEB-1996 (Rel. 33, Created)
DT      01-FEB-1996 (Rel. 33, Last sequence update)
DT      15-MAR-2004 (Rel. 43, Last annotation update)
DE      Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE      (HAF) (Fragment).
GN      F12.
OS      Bos taurus (Bovine).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC      Bovidae; Bovinae; Bos.
OX      NCBI_TaxID=9913;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Liver;
RX      MEDLINE=94242782; PubMed=8186251;
RA      Shibuya Y., Semba U., Okabe H., Kambara T., Yamamoto T.;
RT      "Primary structure of bovine Hageman factor (blood coagulation factor
RT      XII): comparison with human and guinea pig molecules.";
RL      Biochim. Biophys. Acta 1206:63-70(1994).
RN      [2]
RP      SEQUENCE OF 10-21; 350-364 AND 525-550.
RX      MEDLINE=77182112; PubMed=861210;
RA      Fujikawa K., Walsh A.K., Davie W.E.;

```

"Isolation and characterization of bovine factor XII (Hageman factor).";  
 Biochemistry 16:2270-2278 (1977).  
 -|- FUNCTION: Factor XII is a serum glycoprotein that participates in the initiation of blood coagulation, fibrinolysis, and the generation of bradykinin and angiotensin.  
 -|- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor VII to form factor VIIa and factor XI to form factor XIa.  
 -|- PTM: O- AND N-GLYCOSYLATED (BY SIMILARITY).  
 -|- MISCELLANEOUS: Factor XII, prekallikrein, and HWW kininogen form a complex bound to an anionic surface. Prekallikrein is cleaved by factor XII to form kallikrein, which then cleaves factor XII first to alpha-factor XIIa and then to beta-factor XIIa. Alpha-factor XIIa activates factor XI to factor XIa. Bovine factor XII is cleaved only to alpha-factor XIIa as it lacks the trypsin/kallikrein cleavage site.  
 -|- SIMILARITY: Belongs to peptidase family S1.  
 -|- SIMILARITY: Contains 2 EGF-like domains.  
 -|- SIMILARITY: Contains 1 fibronectin type I domain.  
 -|- SIMILARITY: Contains 1 fibronectin type II domain.  
 -|- SIMILARITY: Contains 1 kringle domain.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL: S70184; AABJ0804.2; -;  
 PIR: S45281; S45281.  
 HSSP: P00763; LDPO.  
 MEROPS: S01.211; -;  
 InterPro: IPR009003; Cys\_Ser\_trypsin.  
 InterPro: IPR006209; EGF\_like.  
 InterPro: IPR000083; Fibronctn.  
 InterPro: IPR000562; FN\_Type\_II.  
 InterPro: IPR006210; IEFG.  
 InterPro: IPR000001; Kringle.  
 InterPro: IPR001254; Peptidase\_S1.  
 InterPro: IPR001314; Peptidase\_S1A.  
 Pfam: PF00008; EGF; 2.  
 Pfam: PF00039; fn1; 1.  
 Pfam: PF00040; fn2; 1.  
 Pfam: PF00051; Kringle; 1.  
 Pfam: PF00089; trypsin; 1.  
 PRINTS: PR00722; CHYMOTRYPSIN.  
 PRINTS: PR00013; FNYPEL.  
 PRINTS: PR00018; KRINGLE.  
 ProDom: PD000995; FN\_Type\_II; 1.  
 ProDom: PD000395; Kringle; 1.  
 SMART: SM00181; EGF; 2.  
 SMART: SM00058; FN1; 1.  
 SMART: SM00059; FN2; 1.  
 SMART: SM00130; KR; 1.  
 SMART: SM00020; Tryp\_Spc; 1.  
 PROSITE: PS00022; EGF\_1; 2.  
 PROSITE: PS01186; EGF\_2; FALSE\_NEG.  
 PROSITE: PS50026; EGF\_3; 2.  
 PROSITE: PS01253; FIBRONECTIN\_1; 1.  
 PROSITE: PS00023; FIBRONECTIN\_2; 1.  
 PROSITE: PS00021; KRINGLE\_1; 1.  
 PROSITE: PS50070; KRINGLE\_2; 1.  
 PROSITE: PS50240; TRYPSIN\_DOM; 1.  
 PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 PROSITE: PS00135; TRYPSIN\_SER; 1.  
 KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;  
 KW Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.  
 FT NOW TER 1 9  
 FT SIGNAL <1 9  
 FT CHAIN 10 349  
 FT CHAIN 350 593  
 FT CHAIN ALPHA-FACTOR XIIA HEAVY CHAIN.  
 FT CHAIN ALPHA-FACTOR XIIA LIGHT CHAIN.

FT	DOMAIN	37	78	FIBRONECTIN TYPE-II.
FT	DOMAIN	84	121	EGF-LIKE 1.
FT	DOMAIN	123	163	FIBRONECTIN TYPE-I.
FT	DOMAIN	164	200	EGF-LIKE 2.
FT	DOMAIN	207	287	KRINGLE.
FT	DOMAIN	297	333	PRO-RICH.
FT	DOMAIN	350	593	SERINE PROTEASE.
FT	ACT_SITE	389	389	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	438	438	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	541	541	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	DISULFID	88	100	BY SIMILARITY.
FT	DISULFID	94	109	BY SIMILARITY.
FT	DISULFID	111	120	BY SIMILARITY.
FT	DISULFID	125	153	BY SIMILARITY.
FT	DISULFID	151	160	BY SIMILARITY.
FT	DISULFID	168	179	BY SIMILARITY.
FT	DISULFID	173	188	BY SIMILARITY.
FT	DISULFID	190	199	BY SIMILARITY.
FT	DISULFID	207	287	BY SIMILARITY.
FT	DISULFID	230	269	BY SIMILARITY.
FT	DISULFID	258	282	BY SIMILARITY.
FT	DISULFID	336	463	BY SIMILARITY.
FT	DISULFID	374	390	BY SIMILARITY.
FT	DISULFID	382	452	BY SIMILARITY.
FT	DISULFID	413	416	BY SIMILARITY.
FT	DISULFID	479	547	BY SIMILARITY.
FT	DISULFID	510	526	BY SIMILARITY.
FT	DISULFID	537	568	BY SIMILARITY.
FT	CARBOHYD	99	99	O-LINKED (FUC) (BY SIMILARITY).
FT	CARBOHYD	241	241	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	263	263	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	410	410	N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE	593 AA;	65148 MW;	721592BA792BD61F CRC64;

Query Match 48.3%; Score 43; DB 1; Length 593;  
 Best Local Similarity 41.2%; Pred. No. 6;  
 Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY	1	GMTFRAQEGCAFLTGAAC	17
DB	214	GLSYRGMACTTLTGAPC	230

RESULT 11  
 BMBP BORBU  
 ID BMBP BORBU STANDARD; PRT; 341 AA.  
 AC Q45011; O07954; O31317; O50168; Q44859;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Basic membrane protein B precursor.  
 GN BMBP OR B80382.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sh-2-82;  
 RX MEDLINE=94327086; PubMed=8050720;  
 RA Simpson W.J., Cieplak W., Schrupf M.E., Barbour A.G., Schwan T.G.;  
 RT "Nucleotide sequence and analysis of the gene in Borrelia burgdorferi encoding the immunogenic p39 antigen.";  
 RL FEMS Microbiol. Lett. 119:381-388(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=297;  
 RX MEDLINE=97132632; PubMed=8978084;  
 RA Aron L., Toth C., Godfrey H.P., Cabello P.C.;  
 RT "Identification and mapping of a chromosomal gene cluster of Borrelia burgdorferi containing genes expressed in vivo.";  
 RL FEMS Microbiol. Lett. 145:309-314(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.



SEQUENCE FROM N.A.  
 RC STRAIN=CV. SRI; TISSUE=Leaf;  
 RA Maier T., Zhou L., Thornburg R.W.;  
 RT "Nucleotide sequence of a cDNA encoding UMP synthase from Nicotiana  
 tabacum".  
 RL (In) Plant Gene Register PCR95-025.  
 CC -!- CATALYTIC ACTIVITY: Orotidine 5'-phosphate + diphosphate = orotate  
 CC + 5-phospho-alpha-D-ribose 1-diphosphate.  
 CC -!- CATALYTIC ACTIVITY: Orotidine 5'-phosphate = UMP + CO(2).  
 CC -!- PATHWAY: Pyrimidine biosynthesis; fifth and sixth (last) steps.  
 CC -!- SMILARITY: In the N-terminal section; belongs to the  
 CC purine/pyrimidine phosphoribosyltransferase family.  
 CC -!- SMILARITY: In the C-terminal section; belongs to the OMP  
 CC decarboxylase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@sib-sib.ch).  
 CC  
 CC EMBL; U22260; AAC49115.1; -  
 CC PIR; T02058; T02058.  
 CC HSP; P03962; IDQM.  
 CC InterPro; IPR001754; OMPdecase.  
 CC InterPro; IPR004467; Or pphospho trans.  
 CC InterPro; IPR002375; Pr/pv ip trans.  
 CC InterPro; IPR000836; Prtransferase.  
 CC Pfam; PF00215; OMPdecase; 1.  
 CC Pfam; PF00156; Pribovitran; 1.  
 CC TIGRfams; TIGR00336; Pyre; 1.  
 CC PROSITE; PS00103; PUR PYR PR TRANSFER; 1.  
 CC PROSITE; PS00156; OMPDECASE; 1.  
 CC Pyrimidine biosynthesis; Multifunctional enzyme; Transferase;  
 CC Glycosyltransferase; Lyase; Decarboxylase.  
 CC NON\_TER 1  
 CC ACT\_SITE 291 291 BY SIMILARITY.  
 CC SEQUENCE 461 AA; 49760 MW; AA96479701957CD1 CRC64;  
 Query Match 47.2%; Score 42; DB 1; Length 461;  
 Best Local Similarity 60.0%; Pred. No. 7.1;  
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 TFAAEGAGLTGAAC 17  
 Db 88 TAKATEGAFKPGQAC 102  
 RESULT 14  
 ID DP3A BACHD STANDARD; PRT; 1116 AA.  
 AC Q9K638;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE DNA polymerase III alpha subunit (EC 2.7.7.7).  
 GN DNAE OR BH3169.  
 OS Bacillus halodurans.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 CC NCBI\_TaxID=86665;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C-125 / JCM 9153;  
 RA MEDLINE=20512582; PubMed=11058132;  
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 RA Fuji F., Hirana C., Nakamura Y., Ogasawara N., Kuhara S.,  
 RA Horikoshi K.;  
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";  
 RL Nucleic Acids Res. 28:4317-4331(2000).  
 CC -!- FUNCTION: DNA polymerase III is a complex, multichain enzyme

responsible for most of the replicative synthesis in bacteria.  
 This DNA polymerase also exhibits 3' to 5' exonuclease activity.  
 The alpha chain is the DNA polymerase (By similarity).  
 -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate  
 + [DNA] (N).  
 -!- SUBUNIT: DNA polymerase III contains a core (composed of alpha,  
 epsilon and theta chains) that associates with a tau subunit. This  
 core dimerizes to form the PolIII' complex. PolIII' associates  
 with the gamma complex (composed of gamma, delta, delta', psi and  
 chi chains) and with the beta chain to form the complete DNA  
 polymerase III complex (By similarity).  
 -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 -!- SIMILARITY: Belongs to the DNA polymerase type-C family. DnaE  
 subfamily.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@sib-sib.ch).  
 CC  
 CC EMBL; AP001517; BAB06888.1; -  
 CC PIR; A84046; A84046.  
 CC InterPro; IPR008994; Nucleic\_acid OB.  
 CC InterPro; IPR003141; PHP\_N.  
 CC InterPro; IPR004805; PolC\_alpha.  
 CC InterPro; IPR004365; tRNA\_anti.  
 CC Pfam; PF02231; PHP\_N; 1.  
 CC Pfam; PF01336; tRNA\_anti; 1.  
 CC SMART; SM00481; POLIITAC; 1.  
 CC TIGRfams; TIGR00594; polc; 1.  
 CC Transferase; DNA-directed DNA polymerase; DNA replication;  
 CC Complete proteome.  
 CC SEQUENCE 1116 AA; 127531 MW; 97181EE0BFFD15F1 CRC64;  
 Query Match 47.2%; Score 42; DB 1; Length 1116;  
 Best Local Similarity 60.0%; Pred. No. 17;  
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 GMTFRAQEGAGLTGA 15  
 Db 685 GEARQOEAEAFVTGA 699  
 RESULT 15  
 ID FTSE\_ECOLI STANDARD; PRT; 222 AA.  
 AC P10115;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cell division ATP-binding protein ftse.  
 GN FTSE OR B3463 OR C4256 OR Z4837 OR ECS4312 OR SF3481 OR S4282.  
 OS Escherichia coli O6.  
 OS Escherichia coli O157:H7, and  
 OS Shigella flexneri.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 CC NCBI\_TaxID=562, 217992, 83334, 623;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=E.coli; STRAIN=K12;  
 RX MEDLINE=87089083; PubMed=3025556;  
 RA Gill D.R., Hatfull G.F., Salmond G.P.C.;  
 RT "A new cell division operon in Escherichia coli.";  
 RL Mol. Gen. Genet. 205:134-145(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=E.coli; STRAIN=K12 / MG1655;  
 RX MEDLINE=94316500; PubMed=8041620;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 20:42:21 ; Search time 39 Seconds  
(without alignments)  
137.534 Million cell updates/sec

Title: US-09-982-2598-7

Perfect score: 89

Sequence: 1 GMTFRAOEGAFITGAAC 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL 25:\*
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phase:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_virus:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_unclassified:\*
  - 15: sp\_virus:\*
  - 16: sp\_bacterioph:\*
  - 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72	80.9	243	2	031361
2	72	80.9	264	2	031359
3	72	80.9	325	2	09R776
4	72	80.9	325	2	09R777
5	72	80.9	339	2	087960
6	72	80.9	339	2	0986C1
7	72	80.9	339	2	09PD92
8	69	77.5	264	2	031358
9	45	50.6	374	16	Q8EL43
10	45	50.6	500	12	Q68979
11	45	50.6	626	12	Q9GNF4
12	45	50.6	626	12	Q9GNF5
13	45	50.6	626	12	Q68978
14	45	50.6	626	12	Q9GNF6
15	44	49.4	355	16	Q81WP4
16	44	49.4	361	16	Q81A05

17	43.5	48.9	4928	2	Q9ALM3	Q9alm3 saccharopol
18	43	48.3	256	17	Q9YAR6	Q9yar6 aeropyrum p
19	43	48.3	359	16	Q898S3	Q898s3 clostridium
20	43	48.3	525	16	Q9PORS	Q9pors ureaplasma
21	43	48.3	584	2	Q8RLX0	Q8rlx0 mycoplasma
22	43	48.3	588	2	Q8RLX8	Q8rlx8 mycoplasma
23	42	47.2	210	10	Q8H8C0	Q8h8c0 oryza sativ
24	42	47.2	225	16	Q98CQ2	Q98cq2 rhizobium l
25	42	47.2	239	2	Q8VM53	Q8vm53 salmonella
26	42	47.2	337	2	Q8KWS0	Q8kws0 borrelia bu
27	42	47.2	351	3	Q59701	Q59701 schizosach
28	42	47.2	359	16	Q9WAV7	Q9wv7 thermotoga
29	42	47.2	477	10	Q9LKI3	Q9lki3 nicoitiana p
30	42	47.2	597	11	Q35727	Q35727 mus musculu
31	42	47.2	609	11	Q80YC5	Q80yc5 mus musculu
32	42	47.2	1440	16	Q8XRI6	Q8xri6 ralstonia s
33	41	46.1	188	17	Q8ZZH3	Q8zzh3 pyrobaculum
34	41	46.1	194	16	Q9HU69	Q9hu69 pseudomonas
35	41	46.1	222	2	Q9R630	Q9rf630 escherichia
36	41	46.1	247	2	Q9ZFP3	Q9zf93 bacillus me
37	41	46.1	284	16	Q8A5T6	Q8as5t6 bacteroides
38	41	46.1	347	16	Q8KDL1	Q8kdl1 chlorobium
39	41	46.1	357	16	Q97L60	Q97l60 clostridium
40	41	46.1	375	16	Q82QF4	Q82qf4 streptomyce
41	41	46.1	557	13	Q90XC1	Q90xc1 melagris g
42	41	46.1	563	10	Q8L5K5	Q8l5k5 gossypium h
43	41	46.1	1366	2	Q9L4W5	Q9l4w5 streptomyce
44	40	44.9	148	16	Q88NB8	Q88nb8 pseudomonas
45	40	44.9	196	16	Q8DZ36	Q8dz36 streptococc

## ALIGNMENTS

RESULT 1

031361 PRELIMINARY; PRT; 243 AA.

AC 031361;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Membrane protein A (Fragment).  
GN BMPA.  
OS Borrelia garinii.  
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=29519;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TN;  
RX MEDLINE=98010210; PubMed=9350727;  
RA Roessler D.; Hauser U., Wilske B.;  
RT "Heterogeneity of BmpA (P39) among European Isolates of Borrelia burgdorferi Ssensu Lato and Influence of Interspecies Variability on Serodiagnosis.";  
RL J. Clin. Microbiol. 35:2752-2758(1997).  
DR EMBL: X97243; CAA5882.1;  
DR GO: GO:0008289; F:Filipid binding; IEA.  
DR InterPro; IPR003760; Bmp.  
DR Pfam; PF02608; Bmp; 1.  
FT NON\_TER 1 1  
FT NON\_TER 243 243  
SQ SEQUENCE 243 AA; 26070 MW; 0E750E5E58D0D6F8 CRC64;

Query Match 80.9%; Score 72; DB 2; Length 243;  
Best Local Similarity 100.0%; Pred. No. 0.00025;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMTFRAOEGAFITG 14

Db 115 GMTFRAOEGAFITG 128

RESULT 2

```

O31359
ID O31359 PRELIMINARY; PRT; 264 AA.
AC O31359;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Membrane protein A (Fragment).
DN BmpA.
OS Borrelia garinii.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=29519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=pki;
RX MEDLINE=98010210; PubMed=9350727;
RA Roessler D., Hauser U., Wilske B.;
RT "Heterogeneity of BmpA (P39) among European Isolates of Borrelia burgdorferi Sensu Lato and Influence of Interspecies Variability on Serodiagnosis.";
RL J. Clin. Microbiol. 35:2752-2758(1997).
DR EMBL; X97236; CAA65875.1; -
DR GO; GO:0008289; F.lipid binding; IEA.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
FT NON TER 1
FT 264
SQ SEQUENCE 264 AA; 26284 MW; 5493F14612E127D4 CRC64;

Query Match 80.9%; Score 72; DB 2; Length 264;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFITG 14
Db 115 GMTFRAQEGAFITG 128

RESULT 3
Q9R776
ID Q9R776 PRELIMINARY; PRT; 325 AA.
AC Q9R776;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Membrane protein A (Fragment).
DN BmpA.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T255;
RX MEDLINE=98010210; PubMed=9350727;
RA Roessler D., Hauser U., Wilske B.;
RT "Heterogeneity of BmpA (P39) among European Isolates of Borrelia burgdorferi Sensu Lato and Influence of Interspecies Variability on Serodiagnosis.";
RL J. Clin. Microbiol. 35:2752-2758(1997).
DR EMBL; X97240; CAA65879.1; -
DR GO; GO:0008289; F.lipid binding; IEA.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
FT NON TER 1
FT 325
SQ SEQUENCE 325 AA; 35402 MW; 2F1D43E151CB0663 CRC64;

Query Match 80.9%; Score 72; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFITG 14
Db 115 GMTFRAQEGAFITG 128

RESULT 4
Q9R777
ID Q9R777 PRELIMINARY; PRT; 325 AA.
AC Q9R777;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Membrane protein A (Fragment).
DN BmpA.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PBre;
RX MEDLINE=98010210; PubMed=9350727;
RA Roessler D., Hauser U., Wilske B.;
RT "Heterogeneity of BmpA (P39) among European Isolates of Borrelia burgdorferi Sensu Lato and Influence of Interspecies Variability on Serodiagnosis.";
RL J. Clin. Microbiol. 35:2752-2758(1997).
DR EMBL; X97235; CAA65874.1; -
DR GO; GO:0008289; F.lipid binding; IEA.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
FT NON TER 1
FT 325
SQ SEQUENCE 325 AA; 35436 MW; DF1743E151C10664 CRC64;

Query Match 80.9%; Score 72; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFITG 14
Db 115 GMTFRAQEGAFITG 128

RESULT 5
O87960
ID O87960 PRELIMINARY; PRT; 339 AA.
AC O87960;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE BmpA (p39, ORF1) protein.
DN BmpA (P39, ORF1).
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B31;
RX MEDLINE=98010210; PubMed=9350727;
RA Roessler D., Hauser U., Wilske B.;
RT "Heterogeneity of BmpA (P39) among European Isolates of Borrelia burgdorferi Sensu Lato and Influence of Interspecies Variability on Serodiagnosis.";
RL J. Clin. Microbiol. 35:2752-2758(1997).
DR EMBL; X81515; CAA57235.1; -
DR GO; GO:0008289; F.lipid binding; IEA.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
FT NON TER 1
FT 339
SQ SEQUENCE 339 AA; 37024 MW; D64361DE0CA6C972 CRC64;

Query Match 80.9%; Score 72; DB 2; Length 339;
Best Local Similarity 100.0%; Pred. No. 0.00035;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFITG 14
Db 129 GMTFRAQEGAFITG 142

```



```

RESULT 6
Q9S6C1 PRELIMINARY; PRT; 339 AA.
AC Q9S6C1;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Immunodominant antigen P39.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BT01;
RA Tong Y.P., Feng F.B., Bi S.L., Zhou G.P.;
RT "Molecular cloning and sequencing of P39 gene of a Borrelia
RL burgdorferi strain isolated from China.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF116774; AAD21046.1; -
DR GO; GO:0008289; F:lipid binding; IEA.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
SQ SEQUENCE 339 AA; 36969 MW; A95ABA42EA79596D CRC64;

Query Match 80.9%; Score 72; DB 2; Length 339;
Best Local Similarity 100.0%; Pred. No. 0.00035;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFLTG 14
Db 129 GMTFRAQEGAFLTG 142

RESULT 7
Q9FD92 PRELIMINARY; PRT; 339 AA.
AC Q9FD92;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE BmpA.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N40;
RX MEDLINE=21065112; PubMed=11136458;
RA Dobrikova E.Y., Bugrysheva J., Cabello F.C.;
RT "Two independent transcriptional units control the complex and
RT simultaneous expression of the bmp paralogous chromosomal gene family
RT in Borrelia burgdorferi.";
RL Mol. Microbiol. 39:370-379 (2001).
DR EMBL; AF288609; AAG00584.1; -
DR GO; GO:0008289; F:lipid binding; IEA.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
SQ SEQUENCE 339 AA; 37056 MW; FA90B7F0DE228920 CRC64;

Query Match 80.9%; Score 72; DB 2; Length 339;
Best Local Similarity 100.0%; Pred. No. 0.00035;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFLTG 14
Db 129 GMTFRAQEGAFLTG 142

RESULT 8
O31358 PRELIMINARY; PRT; 264 AA.
ID O31358
AC O31358;

```

```

DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Membrane protein A (Fragment).
GN BmpA.
OS Borrelia garinii.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=29519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PBR;
RX MEDLINE=98010210; PubMed=9350727;
RA Roessler D., Hauser U., Wilske B.;
RT "Heterogeneity of BmpA (P39) among European Isolates of Borrelia
RT burgdorferi Sensu Lato and Influence of Interspecies Variability on
RT Serodiagnosis.";
RL J. Clin. Microbiol. 35:2752-2758 (1997).
DR EMBL; X97242; CAA65881.1; -
DR GO; GO:0008289; F:lipid binding; IEA.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
FT NON TER 1
FT NON TER 264
SQ SEQUENCE 264 AA; 28393 MW; CE0F18C91D0DA95D CRC64;

Query Match 77.5%; Score 69; DB 2; Length 264;
Best Local Similarity 92.9%; Pred. No. 0.0009;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFLTG 14
Db 115 GMTFRAQEGAFLTG 128

```

```

RESULT 9
Q8EL43 PRELIMINARY; PRT; 374 AA.
AC Q8EL43;
DT 01-MAR-2003 (TREMblrel. 23, Created)
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE ABC transporter permease.
GN OB3388.
OS Oceanobacillus iheyensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
OX NCBI_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=22220767; PubMed=12235376;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
RT environments.";
RL Nucleic Acids Res. 30:3927-3935 (2002).
DR EMBL; AP004604; BAC15344.1; -
DR GO; GO:0008289; F:lipid binding; IEA.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Complete proteome.
SQ SEQUENCE 374 AA; 40224 MW; B8680244EE48A48 CRC64;

Query Match 50.6%; Score 45; DB 16; Length 374;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFLTG 16
Db 156 GTFSEHQSGFLVGA 171

```

```

RESULT 10
Q68979 ID Q68979 PRELIMINARY; PRT; 500 AA.
AC Q68979;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 1.9KD ORF.
GN AE.
OS Human herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10298;
RN [1]
RP SEQUENCE OF 1-63 FROM N.A.
RX MEDLINE=84036396; PubMed=6313961;
RA Costa R.H., Draper K.G., Banks L., Powell K.L., Cohen G.,
RA Eisenberg R., Wagner E.K.;
RT "High-resolution characterization of herpes simplex virus type 1
RT transcripts encoding alkaline exonuclease and a 50,000-dalton protein
RT tentatively identified as a capsid protein.";
RL J. Virol. 48:591-603(1983).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86144016; PubMed=3005609;
RA Draper K.G., Devi-Rao G., Costa R.H., Blair E.D., Thompson R.L.,
RA Wagner E.K.;
RT "Characterization of the genes encoding herpes simplex virus type 1
RT and type 2 alkaline exonucleases and overlapping proteins.";
RL J. Virol. 57:1023-1036(1986).
DR EMBL; K02022; AAA45772.1;
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004527; F:exonuclease activity; IEA.
DR InterPro; IPR001616; Herpes_alk_exo.
DR Pfam; PF01771; Herpes_alk_exo; 1.
DR PRINTS; PR00924; ALKXNUCLASE.
SQ SEQUENCE 500 AA; 54412 MW; AES28A06E6FB361E CRC64;

Query Match 50.6%; Score 45; DB 12; Length 500;
Best Local Similarity 56.2%; Pred. No. 28;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFLTGAA 16
Db 411 GVTFRLEDGAGALGAA 426

RESULT 11
Q9QNF4 ID Q9QNF4 PRELIMINARY; PRT; 626 AA.
AC Q9QNF4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Alkaline deoxyribonuclease.
GN UL12.
OS Human herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10298;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99286799; PubMed=10358747;
RA Chiba A., Suzutani T., Saijo M., Koyano S., Azuma M.;
RT "Analysis of nucleotide sequence variations in herpes simplex virus
RT types 1 and 2, and varicella-zoster virus.";
RL Acta Virol. 42:401-407(1998).
DR EMBL; AB009266; BAA84005.2;
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004527; F:exonuclease activity; IEA.
DR InterPro; IPR001616; Herpes_alk_exo.
DR Pfam; PF01771; Herpes_alk_exo; 1.
DR PRINTS; PR00924; ALKXNUCLASE.
SQ SEQUENCE 626 AA; 67492 MW; F5F92710E50A54A2 CRC64;

Query Match 50.6%; Score 45; DB 12; Length 626;
Best Local Similarity 56.2%; Pred. No. 35;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFLTGAA 16
Db 537 GVTFRLEDGAGALGAA 552

RESULT 12
Q9QNF5 ID Q9QNF5 PRELIMINARY; PRT; 626 AA.
AC Q9QNF5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Alkaline deoxyribonuclease.
GN UL12.
OS Human herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10298;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99286799; PubMed=10358747;
RA Chiba A., Suzutani T., Saijo M., Koyano S., Azuma M.;
RT "Analysis of nucleotide sequence variations in herpes simplex virus
RT types 1 and 2, and varicella-zoster virus.";
RL Acta Virol. 42:401-407(1998).
DR EMBL; AB009266; BAA84004.2;
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004527; F:exonuclease activity; IEA.
DR InterPro; IPR001616; Herpes_alk_exo.
DR Pfam; PF01771; Herpes_alk_exo; 1.
DR PRINTS; PR00924; ALKXNUCLASE.
SQ SEQUENCE 626 AA; 67492 MW; F5F92710E50A54A2 CRC64;

Query Match 50.6%; Score 45; DB 12; Length 626;
Best Local Similarity 56.2%; Pred. No. 35;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFLTGAA 16
Db 537 GVTFRLEDGAGALGAA 552

RESULT 13
Q68978 ID Q68978 PRELIMINARY; PRT; 626 AA.
AC Q68978;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Alkaline exonuclease.
GN AE.
OS Human herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10298;
RN [1]
RP SEQUENCE OF 1-189 FROM N.A.
RX MEDLINE=84036396; PubMed=6313961;
RA Costa R.H., Draper K.G., Banks L., Powell K.L., Cohen G.,
RA Eisenberg R., Wagner E.K.;
RT "High-resolution characterization of herpes simplex virus type 1
RT transcripts encoding alkaline exonuclease and a 50,000-dalton protein
RT tentatively identified as a capsid protein.";
RL J. Virol. 48:591-603(1983).
RN [2]

```

```

DR PRINTS; PR00924; ALKXNUCLASE.
SQ SEQUENCE 626 AA; 67484 MW; 88183DEAB9B0DA71 CRC64;

Query Match 50.6%; Score 45; DB 12; Length 626;
Best Local Similarity 56.2%; Pred. No. 35;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFLTGAA 16
Db 537 GVTFRLEDGAGALGAA 552

RESULT 12
Q9QNF5 ID Q9QNF5 PRELIMINARY; PRT; 626 AA.
AC Q9QNF5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Alkaline deoxyribonuclease.
GN UL12.
OS Human herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10298;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99286799; PubMed=10358747;
RA Chiba A., Suzutani T., Saijo M., Koyano S., Azuma M.;
RT "Analysis of nucleotide sequence variations in herpes simplex virus
RT types 1 and 2, and varicella-zoster virus.";
RL Acta Virol. 42:401-407(1998).
DR EMBL; AB009266; BAA84004.2;
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004527; F:exonuclease activity; IEA.
DR InterPro; IPR001616; Herpes_alk_exo.
DR Pfam; PF01771; Herpes_alk_exo; 1.
DR PRINTS; PR00924; ALKXNUCLASE.
SQ SEQUENCE 626 AA; 67492 MW; F5F92710E50A54A2 CRC64;

Query Match 50.6%; Score 45; DB 12; Length 626;
Best Local Similarity 56.2%; Pred. No. 35;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFLTGAA 16
Db 537 GVTFRLEDGAGALGAA 552

RESULT 13
Q68978 ID Q68978 PRELIMINARY; PRT; 626 AA.
AC Q68978;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Alkaline exonuclease.
GN AE.
OS Human herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10298;
RN [1]
RP SEQUENCE OF 1-189 FROM N.A.
RX MEDLINE=84036396; PubMed=6313961;
RA Costa R.H., Draper K.G., Banks L., Powell K.L., Cohen G.,
RA Eisenberg R., Wagner E.K.;
RT "High-resolution characterization of herpes simplex virus type 1
RT transcripts encoding alkaline exonuclease and a 50,000-dalton protein
RT tentatively identified as a capsid protein.";
RL J. Virol. 48:591-603(1983).
RN [2]

```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=86144016; PubMed=3005609;
RA Draper K.G.; Devi-Rao G.; Costa R.H., Blair E.D., Thompson R.L.,
RA Wagner E.K.;
RT "Characterization of the genes encoding herpes simplex virus type 1
RT and type 2 alkaline exonucleases and overlapping proteins.";
RL J. Virol. 57:1023-1036(1986).
DR EMBL; K02022; AAA45771.1;
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004527; F:exonuclease activity; IEA.
DR InterPro; IPR001616; Herpes alk. exo.
DR Pfam; PF01771; Herpes alk. exo. 1.
DR PRINTS; PR00924; ALKEXNUCLASE.
KW Exonuclease.
SQ SEQUENCE 626 AA; 67469 MW; A651D03A29C70260 CRC64;

Query Match 50.6%; Score 45; DB 12; Length 626;
Best Local Similarity 56.2%; Pred. No. 35;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFITGAA 16
Db 537 GVTFRLEDGAGAA 552

RESULT 14
Q9QNF6 PRELIMINARY; PRT; 626 AA.
AC Q9QNF6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Alkaline deoxyribonuclease.
GN UL12.
CS Human herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10298;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=VR-3;
RX MEDLINE=99286799; PubMed=10358747;
RA Chiba A., Suzutani T., Saijo M., Koyano S., Azuma M.;
RT "Analysis of nucleotide sequence variations in herpes simplex virus
RT types 1 and 2, and varicella-zoster virus.";
RL Acta Virol. 42:401-407(1998).
DR EMBL; AB009264; BAA84003.1;
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004527; F:exonuclease activity; IEA.
DR InterPro; IPR001616; Herpes alk. exo.
DR Pfam; PF01771; Herpes alk. exo. 1.
DR PRINTS; PR00924; ALKEXNUCLASE.
SQ SEQUENCE 626 AA; 67558 MW; 68148732098454E7 CRC64;

Query Match 50.6%; Score 45; DB 12; Length 626;
Best Local Similarity 56.2%; Pred. No. 35;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFITGAA 16
Db 537 GVTFRLEDGAGAA 552

RESULT 15
Q81WP4 PRELIMINARY; PRT; 355 AA.
AC Q81WP4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Lipoprotein, Bmp family.
GN BA3927.
OS Bacillus anthracis (strain Ames).
```

```
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=198094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=23608414; PubMed=12721629;
RA Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tetelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Okstad O.A., Helgason E., Rilstone J., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salzberg S.L.,
RA Thomson B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolisto A.-B.,
RA Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria.";
RL Nature 423:81-86(2003).
DR EMBL; AB017036; AAP27658.1;
DR TIGR; BA3927;
DR GO; GO:0008289; F:lipid binding; IEA.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
DR InterPro; IPR000437; Prok_lipoprot_S.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Lipoprotein; Complete proteome.
SQ SEQUENCE 355 AA; 38373 MW; C6E84A9DPE59EF17 CRC64;

Query Match 49.4%; Score 44; DB 16; Length 355;
Best Local Similarity 57.1%; Pred. No. 28;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 MTFRAQEGAFITGAA 15
Db 138 ITFKDHGSGFLVGA 151

Search completed: October 6, 2004, 20:49:04
Job time : 43 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 20:48:11 ; Search time 17 Seconds  
(without alignments)  
51.626 Million cell updates/sec

Title: US-09-982-259B-7

Perfect score: 89  
Sequence: 1 GMTFRAQEGAFITGAAC 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72	80.9	339	1	US-08-396-957A-4
2	43.5	48.9	4928	3	US-09-036-987A-5
3	43.5	48.9	4928	3	US-09-370-700-5
4	43.5	48.9	4928	4	US-09-603-207-5
5	42	47.2	341	1	US-08-396-957A-5
6	41	46.1	111	1	US-08-466-886-43
7	41	46.1	111	3	US-08-469-617-43
8	41	46.1	202	4	US-09-252-991A-33035
9	41	46.1	231	4	US-09-489-039A-10973
10	40	44.9	339	4	US-09-071-035-80
11	40	44.9	361	4	US-09-071-035-78
12	40	44.9	362	4	US-09-134-000C-6004
13	40	44.9	375	4	US-09-134-000C-6005
14	40	44.9	460	2	US-08-677-049-10
15	40	44.9	556	4	US-09-252-991A-27601
16	39.5	44.4	1891	2	US-08-804-227C-12
17	39.5	44.4	1891	2	US-08-804-198-6
18	38.5	43.3	613	4	US-09-331-568A-25
19	38	42.7	749	4	US-09-489-039A-7979
20	37.5	42.1	136	2	US-08-997-080-145
21	37.5	42.1	136	2	US-08-997-362-145
22	37.5	42.1	136	3	US-09-095-855-145
23	37.5	42.1	136	4	US-09-324-542-145
24	37.5	42.1	136	4	US-09-205-426-145
25	37.5	42.1	1482	4	US-09-410-551B-21
26	37.5	42.1	1488	4	US-09-410-551B-17
27	37.5	42.1	1509	4	US-09-410-551B-23

28 37.5 42.1 1517 4 US-09-410-551B-19 Sequence 19, Appl  
29 37.5 42.1 6396 4 US-09-410-551B-72 Sequence 72, Appl  
30 37 41.6 19 4 US-09-431-705-11 Sequence 11, Appl  
31 37 41.6 70 3 US-08-851-843A-190 Sequence 190, App  
32 37 41.6 70 3 US-08-974-549A-309 Sequence 309, App  
33 37 41.6 70 3 US-08-854-050-190 Sequence 190, App  
34 37 41.6 70 4 US-09-430-323-190 Sequence 190, App  
35 37 41.6 70 4 US-09-402-181B-309 Sequence 309, App  
36 37 41.6 70 4 US-09-721-456-309 Sequence 309, App  
37 37 41.6 292 4 US-09-540-236-3287 Sequence 3287, App  
38 37 41.6 330 4 US-09-252-991A-23331 Sequence 23331, A  
39 37 41.6 371 3 US-09-104-308-1 Sequence 1, Appl  
40 37 41.6 371 3 US-09-321-981-1 Sequence 1, Appl  
41 37 41.6 371 4 US-09-739-861A-1 Sequence 1, Appl  
42 37 41.6 371 4 US-09-795-583-1 Sequence 1, Appl  
43 37 41.6 409 3 US-09-140-466-5 Sequence 5, Appl  
44 37 41.6 602 4 US-09-374-454-19 Sequence 19, Appl  
45 37 41.6 1310 4 US-09-170-496D-290 Sequence 290, App

## ALIGNMENTS

RESULT 1  
US-08-396-957A-4  
; Sequence 4, Application US/08396957A  
; Patent No. 5780041  
; GENERAL INFORMATION:  
; APPLICANT: SIMPSON, WARREN; SCHWAN, TOM G.  
; TITLE OF INVENTION: ANTIGENIC PROTEINS AND  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: MORGAN & FINNEGAN, L.L.P.  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WORDPERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/396,957A  
; FILING DATE: 01-MAR-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/020,245  
; FILING DATE: 19-FEB-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/664,731  
; FILING DATE: 05-MAY-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/487,716  
; FILING DATE: 05-MAY-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DOROTHY R. AUTH  
; REGISTRATION NUMBER: 36,434  
; REFERENCE/DOCKET NUMBER: 2026-4018US4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; TELE: 421792  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 339  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein

```

; HYPOTHETICAL: No
; ORIGINAL SOURCE:
; ORGANISM: Borrelia burgdorferi
; STRAIN: Sh-2-82
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAELIOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; FEATURE:
; NAME/KEY: p39
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: p39 protein
; OTHER INFORMATION: sequence.
US-08-396-957A-4

Query Match      80.9%; Score 72; DB 1; Length 339;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GMTFRAQEGAFLTG 14
DB      129 GMTFRAQEGAFLTG 142

RESULT 2
US-09-036-987A-5
; Sequence 5, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Treadway, Patti J.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow Agrosciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R.
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816
; TELEFAX: (317)337-4847
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

Query Match      80.9%; Score 72; DB 1; Length 339;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GMTFRAQEGAFLTG 14
DB      129 GMTFRAQEGAFLTG 142
```

```

US-09-036-987A-5

Query Match      48.9%; Score 43.5; DB 3; Length 4928;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 10; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY      1 GMTFRAQEGAFLTGA 15
DB      3383 GSTY-VREGAFLTGA 3396

RESULT 3
US-09-370-700-5
; Sequence 5, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/370,700
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/36987
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 4928
; TYPE: PRT
; ORGANISM: Saccharopolyspora spinosa
US-09-370-700-5

Query Match      48.9%; Score 43.5; DB 3; Length 4928;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 10; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY      1 GMTFRAQEGAFLTGA 15
DB      3383 GSTY-VREGAFLTGA 3396

RESULT 4
US-09-603-207-5
; Sequence 5, Application US/09603207B
; Patent No. 6521406
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/603,207B
; CURRENT FILING DATE: 2000-06-23
; EARLIER APPLICATION NUMBER: 09/370,700
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 4928
; TYPE: PRT
; ORGANISM: Saccharopolyspora spinosa
US-09-603-207-5

Query Match      48.9%; Score 43.5; DB 4; Length 4928;
```

Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 10; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 GMTFRAQEGAFITGA 15  
Db 3383 GSTY-VREGAFITGA 3396

RESULT 5  
US-08-396-957A-5  
; Sequence 5, Application US/08396957A  
; Patent No. 5780041  
; GENERAL INFORMATION:  
; APPLICANT: SIMPSON, WARREN; SCHWAN, TOM G.  
; TITLE OF INVENTION: ANTIGENIC PROTEINS AND  
; TITLE OF INVENTION: GENES ENCODING SAME OF BORRELIA BURGDORFERI.  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10154

COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WORDPERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/396,957A  
; FILING DATE: 01-MAR-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/020,245  
; FILING DATE: 19-FEB-1993  
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/664,731  
; FILING DATE: 05-MAY-1991  
; APPLICATION NUMBER: 07/487,716  
; FILING DATE: 05-MAY-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DOROTHY R. AUTH  
; REGISTRATION NUMBER: 36,434  
; REFERENCE/DOCKET NUMBER: 2026-4018US4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 751-6849  
; TELEX: 421792  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 341  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: No  
; ORIGINAL SOURCE:  
; ORGANISM: Borrelia burgdorferi  
; STRAIN: Sh-2-82  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; ORGANELLE:  
; FEATURE:  
; NAME/KEY: p39a  
; LOCATION:  
; IDENTIFICATION METHOD:

OTHER INFORMATION: p39a protein sequence  
US-08-396-957A-5

Query Match 47.2%; Score 42; DB 1; Length 341;  
Best Local Similarity 53.8%; Pred. No. 10;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 MTERAQEGAFITG 14  
Db 135 VVFRVQGAFLAG 147

RESULT 6  
US-08-466-886-43  
; Sequence 43, Application US/08466886  
; Patent No. 5776677  
; GENERAL INFORMATION:  
; APPLICANT: Tsui, Lap-Chee  
; APPLICANT: Riordan, John R.  
; APPLICANT: Rommens, Johanna M.  
; APPLICANT: Kerem, Bat-Sheva  
; APPLICANT: Collins, Francis S.  
; APPLICANT: Iannuzzi, Michael C.  
; APPLICANT: Drumm, Mitchell L.  
; APPLICANT: Buckwald, Manuel  
; TITLE OF INVENTION: Cystic Fibrosis Gene  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/466,886  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldstein, Jorge A.  
; REGISTRATION NUMBER: 29,021  
; REFERENCE/DOCKET NUMBER: 1329.0010006  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-2600  
; TELEFAX: 202-371-2540  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 111 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
US-08-466-886-43

Query Match 46.1%; Score 41; DB 1; Length 111;  
Best Local Similarity 62.5%; Pred. No. 4.3;  
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 1 GMTFRAQEG--AFITG 14  
Db 11 GVTFMQFGMAFLTG 26

RESULT 7  
US-08-469-617-43  
; Sequence 43, Application US/08469617  
; Patent No. 6201107  
; GENERAL INFORMATION:

APPLICANT: Tsui, Lap-Chee  
APPLICANT: Riordan, John R.  
APPLICANT: Rommens, Johanna M.  
APPLICANT: Kerem, Bat-Sheva  
APPLICANT: Collins, Francis S.  
APPLICANT: Iannuzzi, Michael C.  
APPLICANT: Drumm, Mitchell L.  
APPLICANT: Buckwald, Manuel  
TITLE OF INVENTION: Cystic Fibrosis Gene  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,617  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldstein, Jorge A.  
REGISTRATION NUMBER: 29,021  
REFERENCE/DOCKET NUMBER: 1329.0010008  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 111 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
US-08-469-617-43

Query Match 46.1%; Score 41; DB 3; Length 111;  
Best Local Similarity 62.5%; Pred. No. 4.3;  
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

Qy 1 GMTFRAQEG--AFLTG 14  
Db 11 GVTFHMQPGEMAFLTG 26

RESULT 8  
US-09-252-991A-33035  
Sequence 33035, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Yaic J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
PRIOR FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 33035  
LENGTH: 202  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-33035

Query Match 46.1%; Score 41; DB 4; Length 202;  
Best Local Similarity 80.0%; Pred. No. 8.6;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 8 EGAFLTGAAC 17  
Db 184 EGREHTGAAC 193

RESULT 9  
US-09-489-039A-10973  
Sequence 10973, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:  
APPLICANT: Gary Breton et. al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
FILE REFERENCE: 2709.2004001  
CURRENT APPLICATION NUMBER: US/09/489,039A  
CURRENT FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: US 60/117,747  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 10973  
LENGTH: 231  
TYPE: PRT  
ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-10973

Query Match 46.1%; Score 41; DB 4; Length 231;  
Best Local Similarity 62.5%; Pred. No. 10;  
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

Qy 1 GMTFRAQEG--AFLTG 14  
Db 29 GVTFHMQPGEMAFLTG 44

RESULT 10  
US-09-071-035-80  
Sequence 80, Application US/09071035  
Patent No. 6448043  
GENERAL INFORMATION:  
APPLICANT: Gil H. Choi  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071,035  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: A. Anders Brookes  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB369P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 80:

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 339 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-071-035-80

Query Match 44.9%; Score 40; DB 4; Length 339;
Best Local Similarity 57.1%; Pred. No. 24;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 TFRAGGAFITGAA 16
Db 126 TFRDNEAAYLAGVA 139

RESULT 11
US-09-071-035-78
; Sequence 78, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 78:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 361 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-071-035-78

Query Match 44.9%; Score 40; DB 4; Length 361;
Best Local Similarity 57.1%; Pred. No. 26;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 TFRAGGAFITGAA 16
Db 148 TFRDNEAAYLAGVA 161

RESULT 12
US-09-134-000C-6004
; Sequence 6004, Application US/09134000C
; Patent No. 6617156
```

```
;
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6004
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-6004

Query Match 44.9%; Score 40; DB 4; Length 362;
Best Local Similarity 57.1%; Pred. No. 26;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 TFRAGGAFITGAA 16
Db 150 TFRDNEAAYLAGVA 163

RESULT 13
US-09-134-000C-6005
; Sequence 6005, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6005
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-6005

Query Match 44.9%; Score 40; DB 4; Length 375;
Best Local Similarity 57.1%; Pred. No. 27;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 TFRAGGAFITGAA 16
Db 162 TFRDNEAAYLAGVA 175

RESULT 14
US-08-677-049-10
; Sequence 10, Application US/08677049
; Patent No. 5858707
; GENERAL INFORMATION:
; APPLICANT: Guimaraes, M. Jorge
; APPLICANT: Bazan, J. Fernando
; APPLICANT: McClanahan, Terrill K.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: PURIFIED MAMMALIAN NUCLEOBASE PERMEASES;
; NUCLEIC ACIDS; ANTIBODIES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
```



COUNTRY: USA  
ZIP: 94304-1104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/677,049  
FILING DATE: 03-JUL-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/000,788  
FILING DATE: 03-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Ching, Edwin P.  
REGISTRATION NUMBER: 34,090  
REFERENCE/DOCKET NUMBER: DX0511  
TELEPHONE: 415-852-9196  
TELEFAX: 415-496-1200  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 460 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Region  
LOCATION: 117..149  
OTHER INFORMATION: /note= "Encompasses TM 4 of Figure"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 328..362  
OTHER INFORMATION: /note= "Encompasses TM 9 of Figure"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 365..392  
OTHER INFORMATION: /note= "Encompasses TM 10 of Figure"  
OTHER INFORMATION: 4"  
US-08-677-049-10

Query Match 44.9%; Score 40; DB 2; Length 460;  
Best Local Similarity 47.1%; Pred. No. 34;  
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAGLTGAC 17  
DB 90 GDTLRGLQALLVSVC 106

RESULT 15  
US-09-252-991A-27601  
Sequence 27601, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 27601  
LENGTH: 556  
TYPE: PRT

ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27601

Query Match 44.9%; Score 40; DB 4; Length 556;  
Best Local Similarity 50.0%; Pred. No. 43;  
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAGLTGAA 16  
DB 288 GQTFRCQPGSSVGSTA 303

Search completed: October 6, 2004, 20:56:48  
Job time: 19 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 20:49:12 ; Search time 128 Seconds  
(without alignments)  
42.739 Million cell updates/sec

Title: US-09-982-259B-7

Perfect score: 89

Sequence: 1 GMTFRAQEGAFITGAAC 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	72	80.9	14	9	US-09-982-264-7
2	72	80.9	14	9	US-09-982-259-7
3	72	80.9	14	10	US-09-982-265-7
4	72	80.9	14	10	US-09-982-287-7
5	46	51.7	1218	16	US-10-437-963-128680
6	45	50.6	626	16	US-10-656-868-2
7	43.5	48.9	4928	16	US-10-323-148A-5
8	42	47.2	1811	16	US-10-437-963-147037
9	41	46.1	102	12	US-10-424-599-231687
10	41	46.1	247	9	US-09-479-040-9
11	41	46.1	375	14	US-10-156-761-8098
12	41	46.1	1366	16	US-10-203-295-5
13	40	44.9	99	12	US-10-424-599-239471
14	40	44.9	117	15	US-10-094-749-2373
15	40	44.9	188	12	US-10-424-599-220111

16	40	44.9	265	14	US-10-017-161-1382	Sequence 1382, Ap
17	40	44.9	328	14	US-10-132-134-23	Sequence 23, Appl
18	40	44.9	339	12	US-10-206-576-80	Sequence 80, Appl
19	40	44.9	361	12	US-10-206-576-78	Sequence 78, Appl
20	40	44.9	449	16	US-10-437-963-148728	Sequence 148728, Ap
21	40	44.9	492	15	US-10-369-493-31	Sequence 31, Appl
22	40	44.9	494	16	US-10-408-765A-1560	Sequence 1560, Ap
23	40	44.9	503	12	US-10-424-599-150986	Sequence 150986, Ap
24	40	44.9	533	12	US-10-424-599-150988	Sequence 233, App
25	40	44.9	604	12	US-10-267-502-233	Sequence 233, App
26	40	44.9	604	16	US-10-287-226-663	Sequence 663, App
27	39	43.8	38	16	US-10-767-701-33126	Sequence 33126, A
28	39	43.8	61	12	US-10-424-599-266236	Sequence 266236, A
29	39	43.8	94	12	US-10-424-599-250461	Sequence 250461, A
30	39	43.8	117	16	US-10-437-963-168300	Sequence 168300, A
31	39	43.8	118	12	US-10-424-599-215876	Sequence 215876, A
32	39	43.8	137	12	US-10-424-599-211728	Sequence 211728, A
33	39	43.8	321	15	US-10-369-493-10365	Sequence 10365, A
34	39	43.8	379	12	US-10-424-599-252634	Sequence 252634, A
35	39	43.8	494	12	US-10-282-122A-67089	Sequence 67089, A
36	39	43.8	814	14	US-10-156-761-7983	Sequence 7983, Ap
37	39	43.8	1256	16	US-10-437-963-130209	Sequence 130209, A
38	38.5	43.3	99	12	US-10-424-599-264860	Sequence 264860, A
39	38.5	43.3	269	12	US-10-424-599-220668	Sequence 220668, A
40	38.5	43.3	613	12	US-10-447-013-25	Sequence 25, Appl
41	38.5	43.3	629	9	US-09-815-242-11063	Sequence 11063, A
42	38.5	43.3	629	12	US-10-282-122A-58229	Sequence 58229, A
43	38	42.7	46	12	US-10-424-599-226124	Sequence 226124, A
44	38	42.7	52	9	US-09-796-692-1393	Sequence 1393, Ap
45	38	42.7	52	14	US-10-040-862-1393	Sequence 1393, Ap

#### ALIGNMENTS

#### RESULT 1

US-09-982-264-7  
; Sequence 7, Application US/09982264  
; Patent No. US20020106706A1  
; GENERAL INFORMATION:  
; APPLICANT: Qiu, Bo  
; APPLICANT: Stein, Stanley  
; APPLICANT: Brunner, Michael  
; APPLICANT: Katz, Michael  
; APPLICANT: Zhang, Guobao  
; APPLICANT: Sigal, Leonard  
; TITLE OF INVENTION: Immunological Test Kit with Borellia burgdorferi Epitope  
; FILE REFERENCE: 271/288  
; CURRENT APPLICATION NUMBER: US/09/982,264  
; CURRENT FILING DATE: 2001-10-17  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patent version 3.1  
; SEQ ID NO 7  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Borellia burgdorferi  
US-09-982-264-7

Query Match 80.9%; Score 72; DB 9; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFITG 14

Db 1 GMTFRAQEGAFITG 14

#### RESULT 2

US-09-982-259-7  
; Sequence 7, Application US/09982259  
; Publication No. US20020197271A1  
; GENERAL INFORMATION:  
; APPLICANT: Qiu, Bo

; APPLICANT: Stein, Stanley  
; APPLICANT: Brunner, Michael  
; APPLICANT: Katz, Michael  
; APPLICANT: Zhang, Guobao  
; APPLICANT: Sigal, Leonard  
; TITLE OF INVENTION: Borellia burgdorferi Epitope Peptides  
; FILE REFERENCE: 271/289  
; CURRENT APPLICATION NUMBER: US/09/982,253  
; CURRENT FILING DATE: 2001-10-17  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Borellia burgdorferi  
US-09-982-259-7

Query Match 80.9%; Score 72; DB 9; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMTFRAQEGAFLTG 14  
Db 1 GMTFRAQEGAFLTG 14

RESULT 3  
US-09-982-265-7  
; Sequence 7, Application US/09982265  
; Publication No. US20030040126A1  
; GENERAL INFORMATION:  
; APPLICANT: Qiu, Bo  
; APPLICANT: Stein, Stanley  
; APPLICANT: Brunner, Michael  
; APPLICANT: Katz, Michael  
; APPLICANT: Zhang, Guobao  
; APPLICANT: Sigal, Leonard  
; TITLE OF INVENTION: Immunological Test Kit with Immunologically Invisible Carrier  
; FILE REFERENCE: 271/287  
; CURRENT APPLICATION NUMBER: US/09/982,265  
; CURRENT FILING DATE: 2001-10-17  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Borellia burgdorferi  
US-09-982-265-7

Query Match 80.9%; Score 72; DB 10; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMTFRAQEGAFLTG 14  
Db 1 GMTFRAQEGAFLTG 14

RESULT 4  
US-09-982-287-7  
; Sequence 7, Application US/09982287  
; Publication No. US20030040127A1  
; GENERAL INFORMATION:  
; APPLICANT: Qiu, Bo  
; APPLICANT: Stein, Stanley  
; APPLICANT: Brunner, Michael  
; APPLICANT: Katz, Michael  
; APPLICANT: Zhang, Guobao  
; APPLICANT: Sigal, Leonard  
; TITLE OF INVENTION: Multiple Epitopes Connected by Carrier  
; FILE REFERENCE: 271/090  
; CURRENT APPLICATION NUMBER: US/09/982,287  
; CURRENT FILING DATE: 2001-10-17

; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Borellia burgdorferi  
US-09-982-287-7

Query Match 80.9%; Score 72; DB 10; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMTFRAQEGAFLTG 14  
Db 1 GMTFRAQEGAFLTG 14

RESULT 5  
US-10-437-963-128680  
; Sequence 128680, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: Ia Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 128680  
; LENGTH: 1218  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_31010C.1.pep  
US-10-437-963-128680

Query Match 51.7%; Score 46; DB 16; Length 1218;  
Best Local Similarity 61.5%; Pred. No. 58;  
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MTFRAQEGAFLTG 14  
Db 1166 VTFRAKKGSLAG 1178

RESULT 6  
US-10-656-868-2  
; Sequence 2, Application US/10656868  
; Publication No. US20040141994A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Connecticut  
; APPLICANT: Weller, Sandra K  
; APPLICANT: Bacher Reuven, Nina  
; APPLICANT: Myers, Richard S  
; TITLE OF INVENTION: Viral Recombinases, Related Articles and Methods of Use Thereof  
; FILE REFERENCE: UCT-0039  
; CURRENT APPLICATION NUMBER: US/10/656,868  
; CURRENT FILING DATE: 2003-09-04  
; PRIOR APPLICATION NUMBER: US 60/408,092  
; PRIOR FILING DATE: 2002-09-04  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2  
; LENGTH: 626  
; TYPE: PRT

ORGANISM: Herpes simplex virus 1  
 PUBLICATION INFORMATION:  
 DATABASE ACCESSION NUMBER: GI 119693  
 DATABASE ENTRY DATE: 1992-05-01  
 RELEVANT RESIDUES: (1)..(626)  
 US-10-656-868-2

Query Match 50.6%; Score 45; DB 16; Length 626;  
 Best Local Similarity 56.2%; Pred. No. 42;  
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFITGAA 16  
 Db 537 GVTFRLEDGAGALGA 552

RESULT 7

US-10-329-148A-5  
 Sequence 5, Application US/10329148A  
 Publication No. US20040023343A1  
 GENERAL INFORMATION:  
 APPLICANT: Baltz, Richard H  
 APPLICANT: Broughton, Mary C  
 APPLICANT: Crawford, Kathryn P  
 APPLICANT: Madduri, Krishnamurthy  
 APPLICANT: Treadway, Patti J  
 APPLICANT: Turner, Jan R  
 APPLICANT: Waldron, Clive

TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide  
 FILE REFERENCE: 50489 DIV1  
 CURRENT APPLICATION NUMBER: US/10/329,148A  
 CURRENT FILING DATE: 2002-12-23  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/603,207B  
 PRIOR FILING DATE: EARLIER FILING DATE: 2000-06-23  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/370,700  
 PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-09  
 NUMBER OF SEQ ID NOS: 39  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 5

LENGTH: 4928

TYPE: PRT  
 ORGANISM: Saccharopolyspora spinosa  
 US-10-329-148A-5

Query Match 48.9%; Score 43.5; DB 16; Length 4928;  
 Best Local Similarity 66.7%; Pred. No. 7.2e+02;  
 Matches 10; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 GMTFRAQEGAFITGA 15  
 Db 3383 GSTY-VREGAFITGA 3396

RESULT 8

US-10-437-963-147037  
 Sequence 147037, Application US/10437963  
 Publication No. US20040123343A1  
 GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 APPLICANT: Wu, Wei  
 APPLICANT: Boukharov, Andrey A.  
 APPLICANT: Barbazuk, Brad  
 APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(53221)B  
 CURRENT APPLICATION NUMBER: US/10/437,963  
 CURRENT FILING DATE: 2003-05-14  
 NUMBER OF SEQ ID NOS: 204966  
 SEQ ID NO 147037

LENGTH: 1811  
 TYPE: PRT  
 ORGANISM: Oryza sativa  
 FEATURE:  
 OTHER INFORMATION: Clone ID: PAT\_MRT4530\_47603C.1.pep  
 US-10-437-963-147037

Query Match 47.2%; Score 42; DB 16; Length 1811;  
 Best Local Similarity 56.2%; Pred. No. 4.4e+02;  
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 GMTFRAQEGAFITGAA 16  
 Db 68 GGAFRDEHGAFAVAGYA 83

RESULT 9

US-10-424-599-231687  
 Sequence 231687, Application US/10424599  
 Publication No. US20040031072A1  
 GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J  
 APPLICANT: Kovalic, David K  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(53223)B  
 CURRENT APPLICATION NUMBER: US/10/424,599  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 285684  
 SEQ ID NO 231687  
 LENGTH: 102  
 TYPE: PRT  
 ORGANISM: Glycine max  
 FEATURE:  
 OTHER INFORMATION: Clone ID: PAT\_MRT3847\_51236C.1.pep  
 US-10-424-599-231687

Query Match 46.1%; Score 41; DB 12; Length 102;  
 Best Local Similarity 50.0%; Pred. No. 29;  
 Matches 8; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 MTFRAGGAFITGAC 17  
 Db 52 MTMALPRGAFITGACC 67

RESULT 10

US-09-479-040-9  
 Sequence 9, Application US/09479040  
 Publication No. US20020182690A1  
 GENERAL INFORMATION:

APPLICANT: McCool, Gabriel J.  
 APPLICANT: Cannon, Maura C.  
 APPLICANT: Cannon, Francis C.  
 APPLICANT: Valentin, Henry E.  
 APPLICANT: Gruys, Kenneth J.

TITLE OF INVENTION: POLYHYDROXYALKANOATE BIOSYNTHESIS ASSOCIATED PROTEINS  
 TITLE OF INVENTION: AND CODING REGION IN BACILLUS MEGATERIUM  
 FILE REFERENCE: MOBT212  
 CURRENT APPLICATION NUMBER: US/09/479,040  
 CURRENT FILING DATE: 2000-01-07  
 NUMBER OF SEQ ID NOS: 29  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 9  
 LENGTH: 247  
 TYPE: PRT  
 ORGANISM: Bacillus megaterium  
 US-09-479-040-9

Query Match 46.1%; Score 41; DB 9; Length 247;  
 Best Local Similarity 42.9%; Pred. No. 75;

Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;  
QY 1 GMTFRAOEGAFLTG 14  
Db 223 GVVYLANDGAYITG 236

RESULT 11  
US-10-156-761-8098  
; Sequence 8098, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; PRIOR FILING DATE: 2002-05-29  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 8098  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-8098

Query Match 46.1%; Score 41; DB 14; Length 375;  
Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 TFAOEGAFLTGAAC 17  
Db 358 TFAAPAGFSLNGAAC 372

RESULT 12  
US-10-203-295-5  
; Sequence 5, Application US/10203295  
; Publication No. US20040115762A1  
; GENERAL INFORMATION:  
; APPLICANT: Zotchev, Sergey Borisovich  
; APPLICANT: Sekurova, Olga Nikalayivna  
; APPLICANT: Fjaervik, Epsen  
; APPLICANT: Brataaset, Trygve  
; APPLICANT: Strom, Arne Reidar  
; APPLICANT: Valla, Svein  
; APPLICANT: Ellingsen, Trond Erling  
; APPLICANT: Sletta, Havard  
; TITLE OF INVENTION: Novel genes encoding a nystatin polyketide synthase and their  
; TITLE OF INVENTION: manipulation and utility  
; FILE REFERENCE: 1181-265  
; CURRENT APPLICATION NUMBER: US/10/203,295  
; CURRENT FILING DATE: 2003-05-19  
; PRIOR APPLICATION NUMBER: PCT/GB 01/00509  
; PRIOR FILING DATE: 2001-02-08  
; PRIOR APPLICATION NUMBER: GB 0002840.7  
; PRIOR FILING DATE: 2000-02-08  
; PRIOR APPLICATION NUMBER: GB 0008786.6  
; PRIOR FILING DATE: 2000-04-10  
; PRIOR APPLICATION NUMBER: GB 0009387.2  
; PRIOR FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 5

; LENGTH: 1366  
; TYPE: PRT  
; ORGANISM: Streptomyces noursei ATCC 11455  
US-10-203-295-5  
Query Match 46.1%; Score 41; DB 16; Length 1366;  
Best Local Similarity 80.0%; Pred. No. 4.8e+02;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 7 OEGAFLTGAA 16  
Db 65 REGGFLTGA 74

RESULT 13  
US-10-424-599-239471  
; Sequence 239471, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 239471  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT MRT3847\_58268C.1.pcp  
US-10-424-599-239471

Query Match 44.9%; Score 40; DB 12; Length 99;  
Best Local Similarity 66.7%; Pred. No. 41;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 MTFRAOEGAFLT 13  
Db 18 MTFIAYEGSVLT 29

RESULT 14  
US-10-094-749-2373  
; Sequence 2373, Application US/10094749  
; Publication No. US20030219741A1  
; GENERAL INFORMATION:  
; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: OTSUKI, TETSUJI  
; APPLICANT: WAKAMATSU, AI  
; APPLICANT: SATO, HIROYUKI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: YAMAMOTO, JUN-ICHI  
; APPLICANT: ISONO, YUUKO  
; APPLICANT: HIO, YURI  
; APPLICANT: OTSUKA, KAORU  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: IRIE, RYOTARO  
; APPLICANT: TAMECHIKA, ICHIRO  
; APPLICANT: SEKI, NAOHICO  
; APPLICANT: YOSHIKAWA, TSUTOMU  
; APPLICANT: OTSUKA, MOTOYUKI  
; APPLICANT: NAGAHARI, KENJI  
; APPLICANT: MASUHO, YASUHIKO  
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA  
; FILE REFERENCE: 084335/0160  
; CURRENT APPLICATION NUMBER: US/10/094,749  
; CURRENT FILING DATE: 2002-03-12

; PRIOR APPLICATION NUMBER: 60/350,435  
 ; PRIOR FILING DATE: 2002-01-24  
 ; PRIOR APPLICATION NUMBER: JP 2001-328381  
 ; PRIOR FILING DATE: 2001-09-14  
 ; NUMBER OF SEQ ID NOS: 3381  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2373  
 ; LENGTH: 117  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-094-749-2373

Query Match 44.9%; Score 40; DB 15; Length 117;  
 Best Local Similarity 40.0%; Pred. No. 50;  
 Matches 10; Conservative 1; Mismatches 6; Indels 8; Gaps 1;

QY 1 GMTFRAQEG-----AFLTGAC 17  
 DB 21 GVTLRTEGWFVRPLFAPCKGACC 45

RESULT 15  
 US-10-424-599-220111  
 ; Sequence 220111, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 220111  
 ; LENGTH: 188  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT MRT3847\_4078C.1.pep  
 ; US-10-424-599-220111

Query Match 44.9%; Score 40; DB 12; Length 188;  
 Best Local Similarity 47.1%; Pred. No. 83;  
 Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 GMTFRAQEGCAFLTGAC 17  
 DB 31 GMIQKSGNGAIPTGRPC 47

Search completed: October 6, 2004, 21:00:21  
 Job time : 131 secs